

**LAPORAN UJIAN AKHIR SEMESTER**  
**STATISTIKA DESKRIPTIF**  
**CORRESPONDENCE ANALYSIS AND MULTIPLE**  
**CORRESPONDENCE ANALYSIS (CA & MCA)**



**NAMA : MUKHAMAD IKHSANUDIN**  
**NIM : 082011633086**  
**DOSEN PENGAMPU : Drs. ETO WURYANTO, DEA.**  
**196609281991021001**

**PROGRAM STUDI S1 SISTEM INFORMASI**  
**FAKULTAS SAINS DAN TEKNOLOGI**  
**UNIVERSITAS AIRLANGGA**

**2021**

```

```{R}
# CA & MCA
library(flexclust)
library(cluster)
library(factoextra)
library(mclust)
library(FactoMineR)
library(corrplot)
library(ade4)
library(MASS)

# CA
# Dataset Preparation
load("Tithonia.Rda")
Data_CA <- Tithonia
head(Data_CA)

```

Description: `df[,14]` [6 x 14]

	Seed Size <dbl>	Flower Size <dbl>	Female Leaf Size <dbl>	Head Size <dbl>	Number of Flowers <int>
T_calva_calva	4.10	5.98	10.62	2.72	140
T_lancifolia	3.77	5.59	7.84	2.35	120
T_longiradiata	3.73	5.97	45.68	2.22	140
T_pedunculata	6.10	7.10	21.52	3.01	100
T_fruticosa	4.57	11.25	46.71	2.49	120
T_koelzii	5.52	4.50	62.16	4.50	100

6 rows | 1-6 of 14 columns

```

my_CA <- CA(Data_CA, graph = FALSE)
my_CA

```

```

**Results of the Correspondence Analysis (CA)**
The row variable has 11 categories; the column variable has 14 categories
The chi square of independence between the two variables is equal to 1191.879 (p-value = 5.566012e-171 ).
*The results are available in the following objects:

  name          description
1  "$eig"        "eigenvalues"
2  "$col"        "results for the columns"
3  "$col$coord"  "coord. for the columns"
4  "$col$cos2"   "cos2 for the columns"
5  "$col$contrib" "contributions of the columns"
6  "$row"        "results for the rows"
7  "$row$coord"  "coord. for the rows"
8  "$row$cos2"   "cos2 for the rows"
9  "$row$contrib" "contributions of the rows"
10 "$call"       "summary called parameters"
11 "$call$marge.col" "weights of the columns"
12 "$call$marge.row" "weights of the rows"

```

CA merupakan fungsi untuk menghitung analisis korespodensi dan mendapatkan hasil sebagai berikut, variabel baris memiliki 27 kategori dan variabel kolom memiliki 3 kategori Hasilnya tersedia di 12 objek.

```

# Statistifical Significance
# Chi-Square Statistics
Chi <- 42.652
# Degree of Freedom
DF1 <- (1/(nrow(Data_CA)-1))*100
DF1
DF2 <- (1/(ncol(Data_CA)-1))*100

```

```

DF2
# P-Value
PVal <- pchisq(Chi, DF1 == DF2, lower.tail = FALSE)
PVal

```

```

[1] 10
[1] 7.692308
[1] 0

```

Chi square bernilai 42,652

Derajat kebebasan untuk baris bernilai 10%

Derajat kebebasan untuk kolom bernilai 7,69%

Nilai P adalah 0

```

# Eigenvalues / Variances
get_eig(my_CA)

```

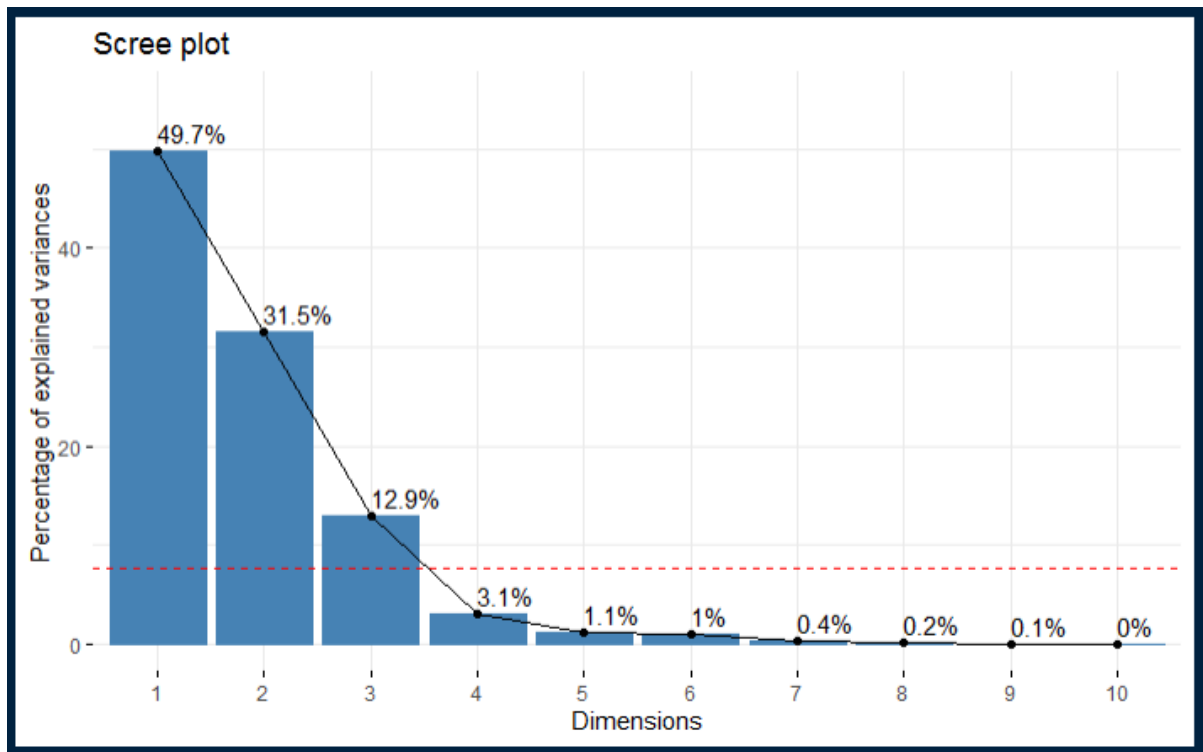
	eigenvalue	variance.percent	cumulative.variance.percent
Dim.1	1.355662e-01	49.74614052	49.74614
Dim.2	8.587035e-02	31.51019881	81.25634
Dim.3	3.513677e-02	12.89346799	94.14981
Dim.4	8.441964e-03	3.09778591	97.24759
Dim.5	3.039781e-03	1.11545024	98.36304
Dim.6	2.594188e-03	0.95193968	99.31498
Dim.7	1.030218e-03	0.37803957	99.69302
Dim.8	6.054775e-04	0.22218051	99.91520
Dim.9	1.541190e-04	0.05655410	99.97176
Dim.10	7.696579e-05	0.02824266	100.00000

Dari perhitungan tersebut didapat 10 nilai eigen yang diurutkan dari varians terbesar, sehingga pada dimensi 1 nilai eigennya adalah 0.0135 dan dimensi berikutnya sampai dimensi 10.

```

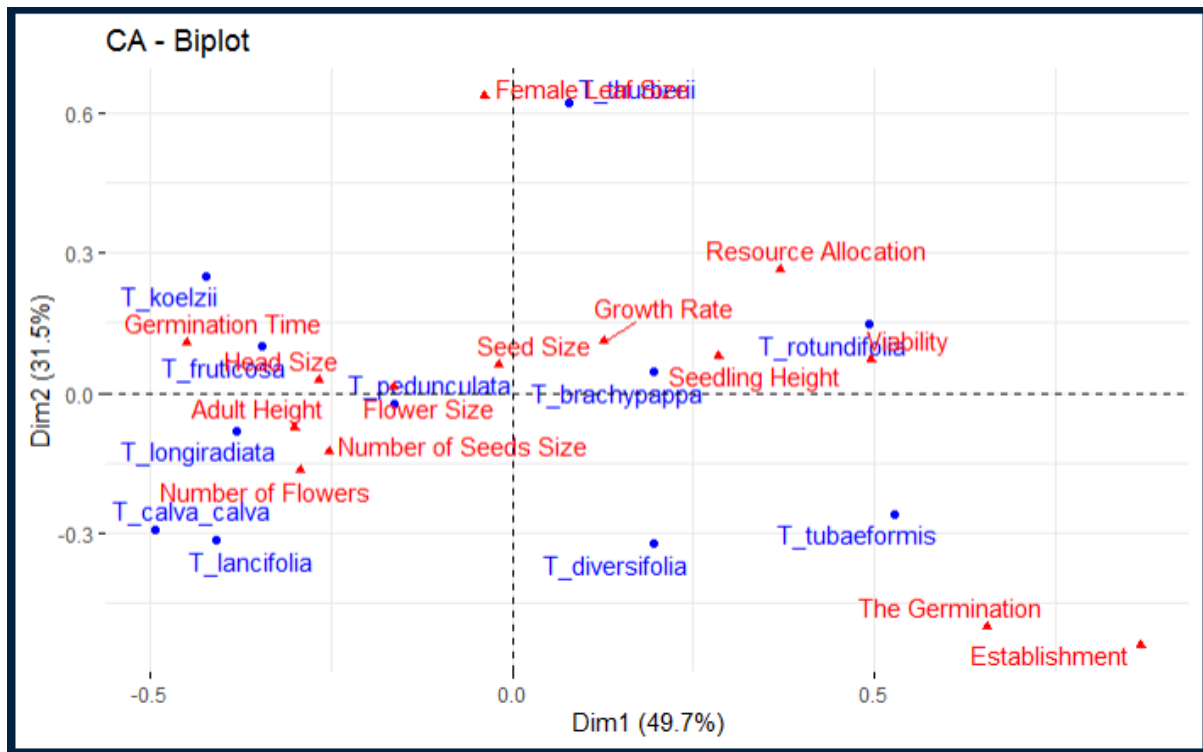
fviz_screplot(my_CA, addlabels = TRUE, ylim = c(0, 55)) +
  geom_hline(yintercept = DF2, linetype = 2, color = "red")

```



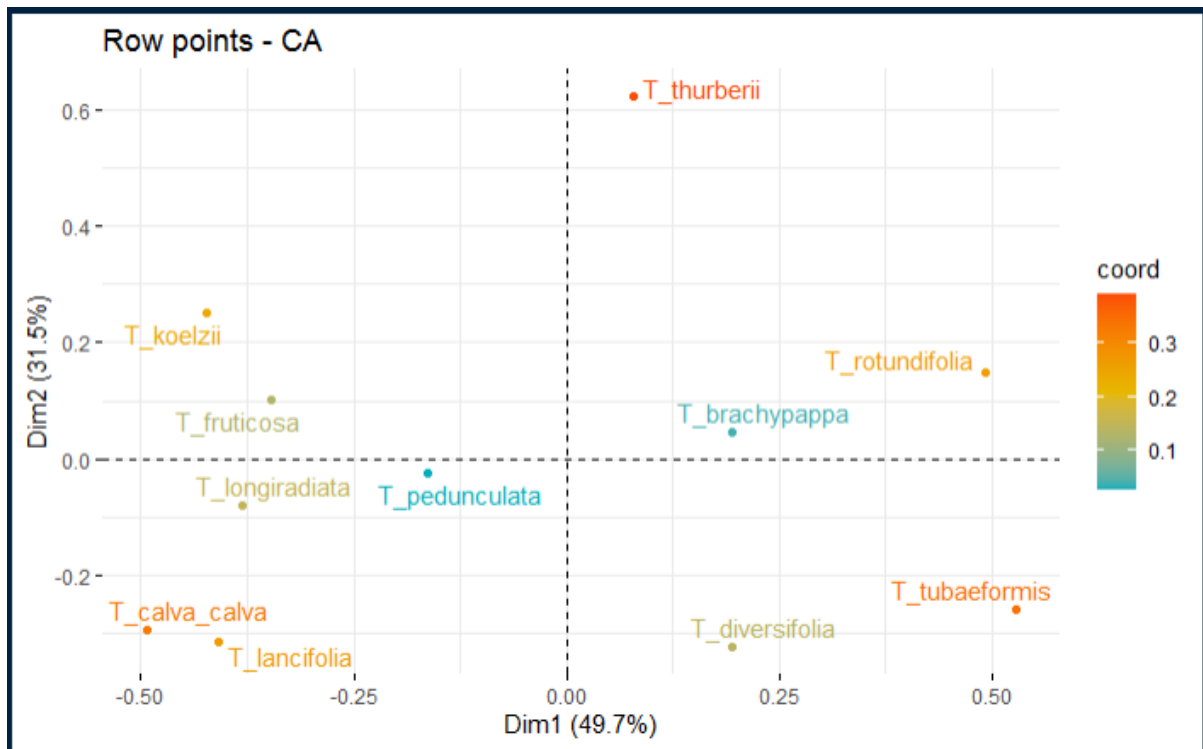
Berdasarkan scree plot pada dimensi 1, distribusi data yang dapat dilakukan analisis berada pada persentase variance 49,7 % dimana pada dimensi 1 ini distribusi data terletak pada kondisi distribusi data yang terbaik. Hal ini karena semakin besar persentase variancinya menunjukkan bahwa data tersebut telah banyak terdistribusi dan masuk untuk dilakukan analisis. Lalu distribusi data pada dimensi 2 dengan variance 31,5% dan seterusnya. Dari plot di atas, terdapat 3 dimensi yang memenuhi derajat kebebasan.

```
# Biplot  
fviz_ca_biplot(my_CA, rep = TRUE)
```



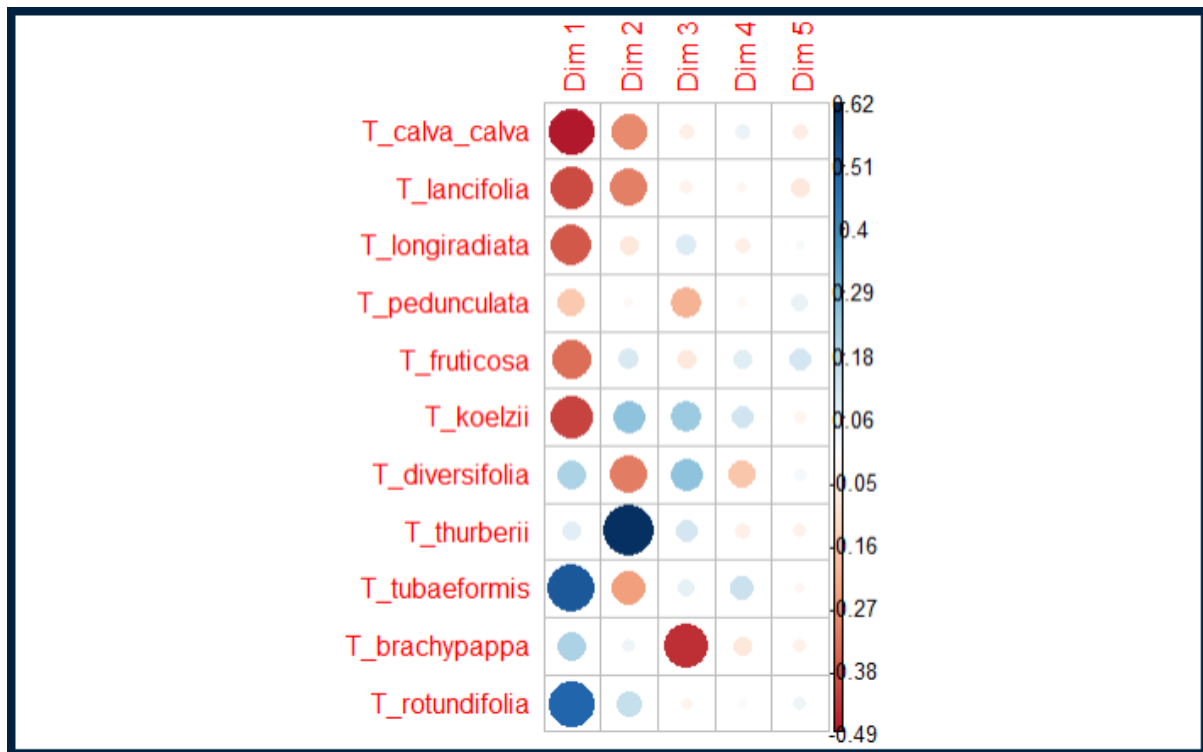
Berdasarkan biplot di atas, terdapat 90,2% persebaran data yang masuk. Dari plot di atas, kita tahu berbagai spesies (biru) dengan karakteristiknya (merah). Terdapat beberapa karakteristik yang sangat dominan bagi suatu spesies. *T\_thuberii* sangat dominan dengan Female Leaf Size nya (memiliki nilai Female Leaf Size terbesar) dan seterusnya.

```
# Graph or Row variables
my_CA_Row <- my_CA$row
# Coord
fviz_ca_row(my_CA, col.row = "coord",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE)
```



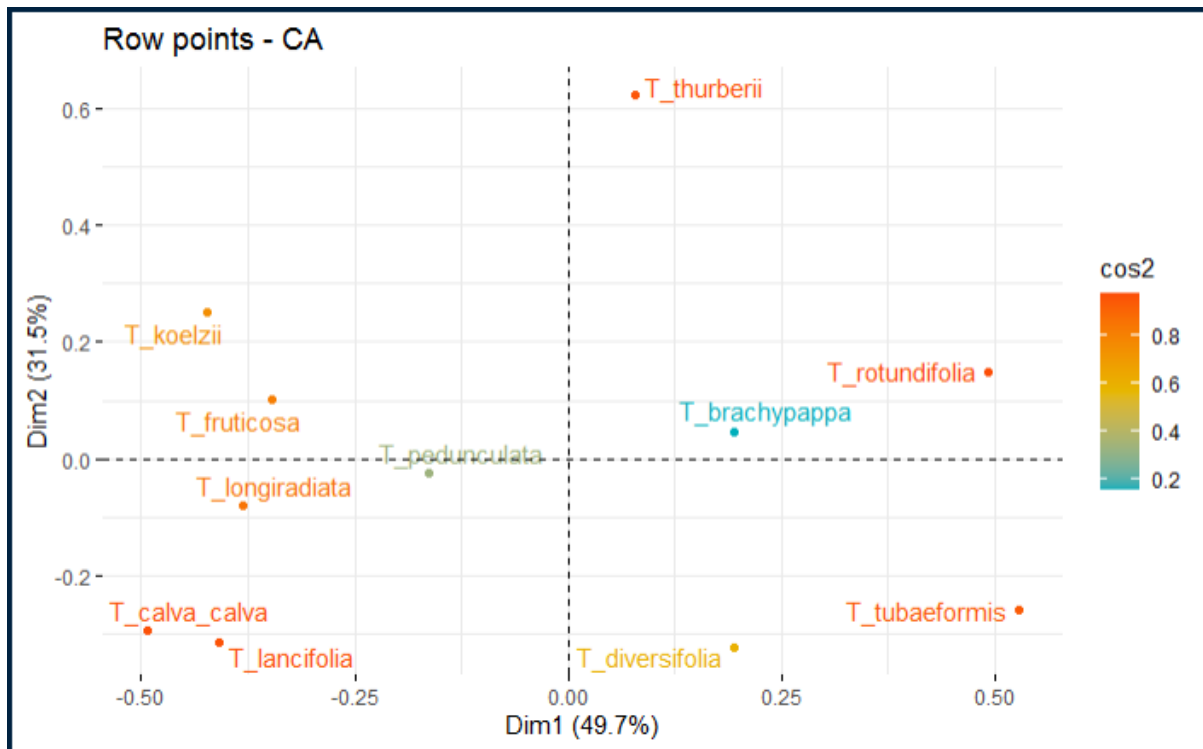
Dari plot CA di atas, data yang telah tersebar adalah 90,2% dan didapati bahwa *T\_thurberii* memiliki nilai positif pada kedua dimensi dan dominan dengan dimensi 2. Sedangkan *T\_rotundifolia* dan *T\_brachypappa* memiliki nilai positif pada keduanya dan dominan dengan dimensi 1. Lalu terdapat *T\_koelzii* dan *T\_fruticosa* yang memiliki nilai dimensi 2 positif dan dimensi 1 negatif. *T\_diversifolia* dan *T\_tubaeformis* memiliki nilai dimensi 1 positif namun nilai dimensi 2 nya negatif. Serta *T\_calva\_calva*, *T\_lancifolia*, *T\_longiradiata* dan *T\_pedunculata* memiliki nilai negatif di kedua dimensinya

```
corrplot(my_CA_Row$coord, is.corr = FALSE)
```



- Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi baris tertinggi hingga terendah. Pada dimensi 1, yang terendah adalah T\_calva\_calva dan yang tertinggi adalah T\_rotundifolia.
- Sedangkan untuk kontribusinya, T\_calva\_calva merupakan yang tertinggi dan T\_pedunculata merupakan yang terendah

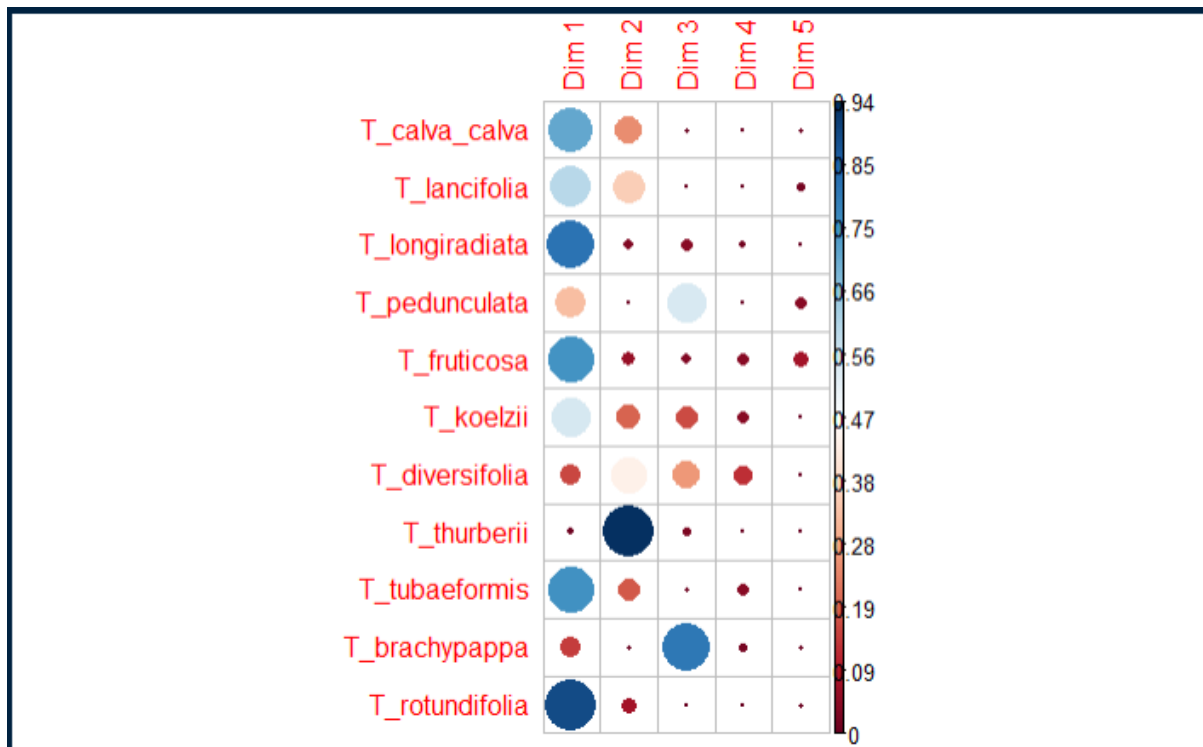
```
# Cos2
fviz_ca_row(my_CA, col.row = "cos2",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE)
```



Dari plot CA di atas, data yang telah tersebar adalah 90,2% dan didapati bahwa *T\_thurberii* memiliki nilai positif pada kedua dimensi dan dominan dengan dimensi 2. Sedangkan *T\_rotundifolia* dan *T\_brachypappa* memiliki nilai positif pada keduanya dan dominan dengan dimensi 1. Lalu terdapat *T\_koelzii* dan *T\_fruticosa* yang memiliki nilai dimensi 2 positif dan dimensi 1 negatif. *T\_diversifolia* dan *T\_tubaeformis* memiliki nilai dimensi 1 positif namun nilai dimensi 2 nya negatif. Serta *T\_calva\_calva*, *T\_lancifolia*, *T\_longiradiata* dan *T\_pedunculata* memiliki nilai negatif di kedua dimensinya. Yang berbeda dari plot CA sebelumnya adalah skala visualisasi warnanya. Pada coord, nilai tertinggi adalah 0.5, sedangkan pada cos2 nilai tertinggi adalah 1.

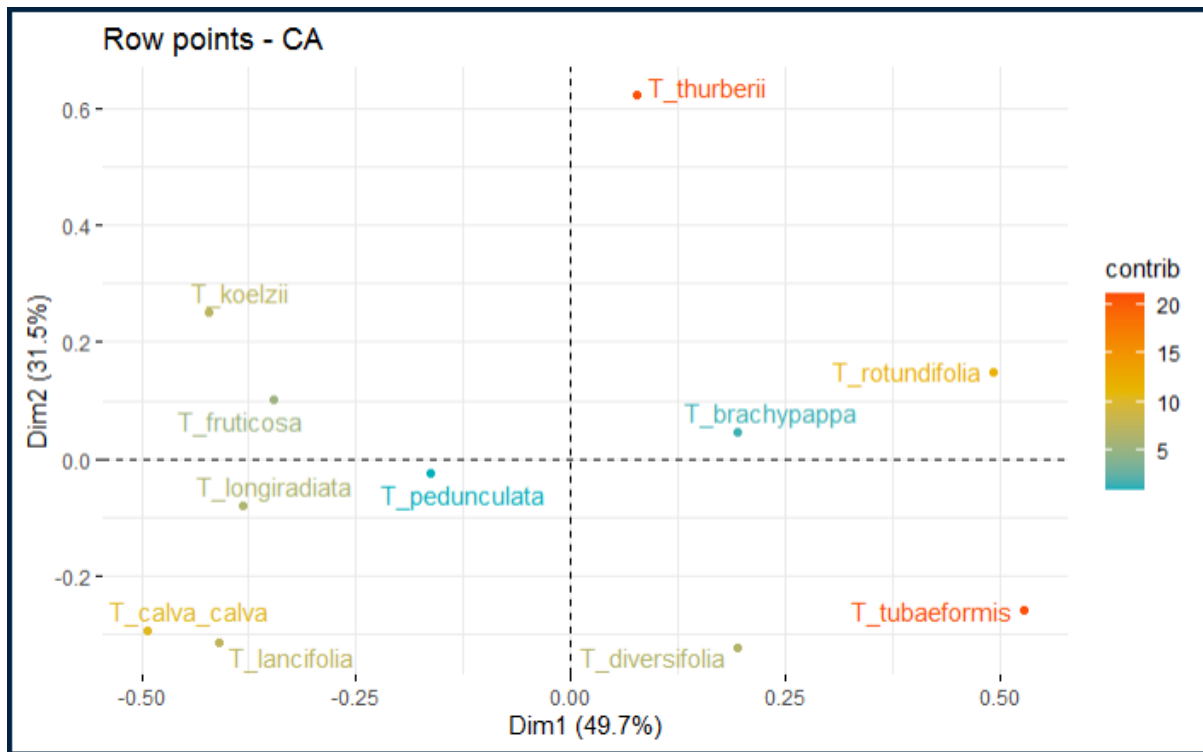
```
corrplot(my_CA_Row$cos2, is.corr = FALSE)
```





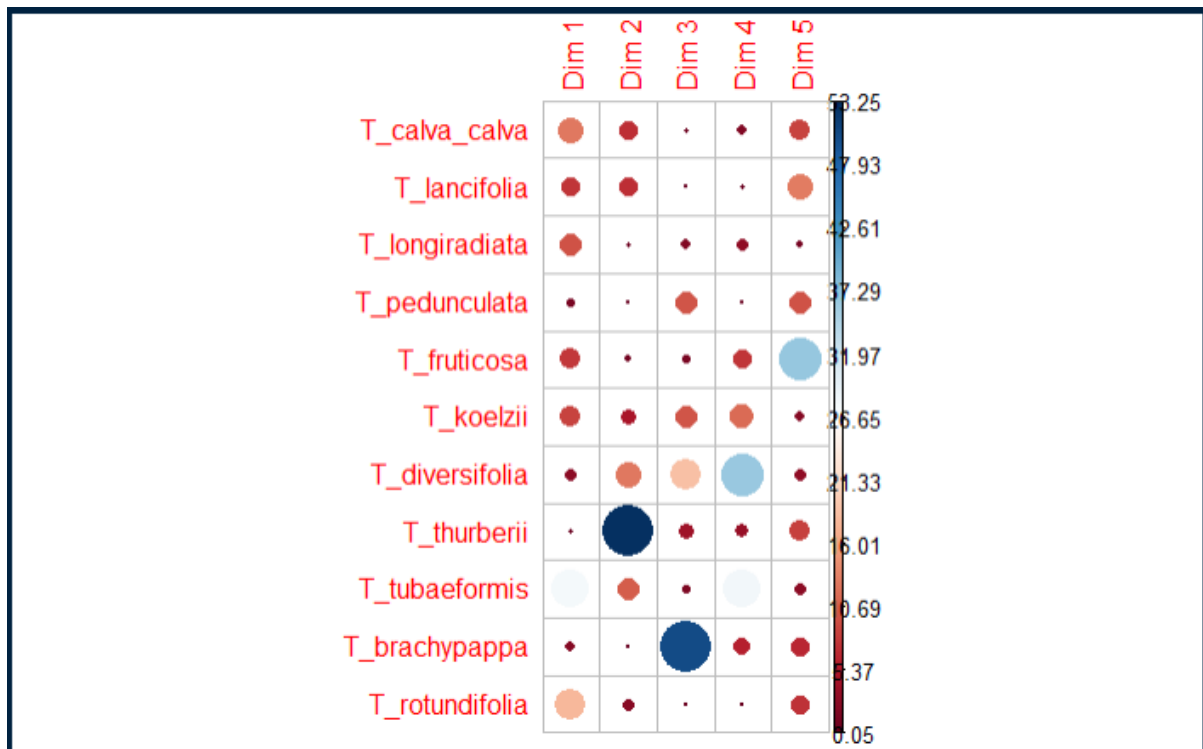
- Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi baris tertinggi hingga terendah. Pada dimensi 1, yang terendah adalah T\_thurberii dan yang tertinggi adalah T\_rotundifolia.
- Sedangkan untuk kontribusinya, T\_lancifolia merupakan yang tertinggi dan T\_diversifolia merupakan yang terendah

```
# Contrib
fviz_ca_row(my_CA, col.row = "contrib",
            gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
            repel = TRUE)
```



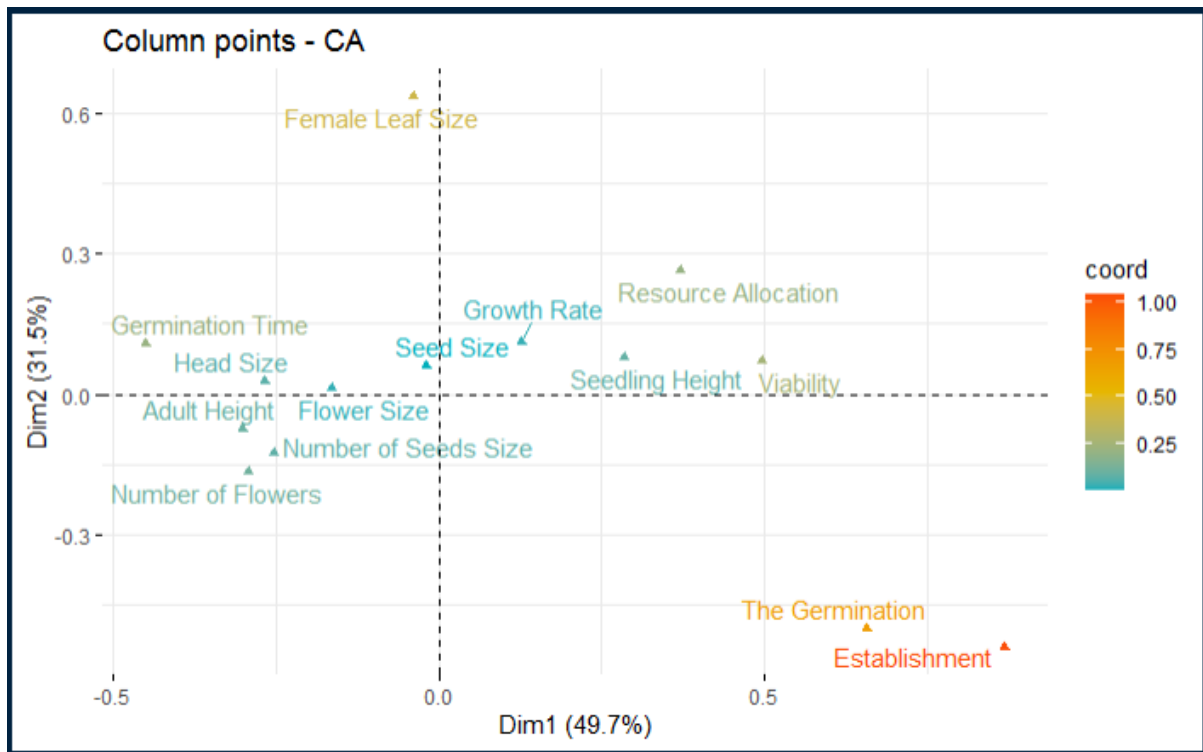
Dari plot CA di atas, data yang telah tersebar adalah 90,2% dan didapatkan bahwa *T\_thurberii* memiliki nilai positif pada kedua dimensi dan dominan dengan dimensi 2. Sedangkan *T\_rotundifolia* dan *T\_brachypappa* memiliki nilai positif pada keduanya dan dominan dengan dimensi 1. Lalu terdapat *T\_koelzii* dan *T\_fruticosa* yang memiliki nilai dimensi 2 positif dan dimensi 1 negatif. *T\_diversifolia* dan *T\_tubaeformis* memiliki nilai dimensi 1 positif namun nilai dimensi 2 nya negatif. Serta *T\_calva\_calva*, *T\_lancifolia*, *T\_longiradiata* dan *T\_pedunculata* memiliki nilai negatif di kedua dimensinya. Yang berbeda dari plot CA sebelumnya adalah skala visualisasi warnanya. Pada coord, nilai tertinggi adalah 0.5, pada  $\cos^2$  nilai tertinggi adalah 1, dan pada contrib, nilai tertinggi adalah 20.

```
corrplot(my_CA_Row$contrib, is.corr = FALSE)
```



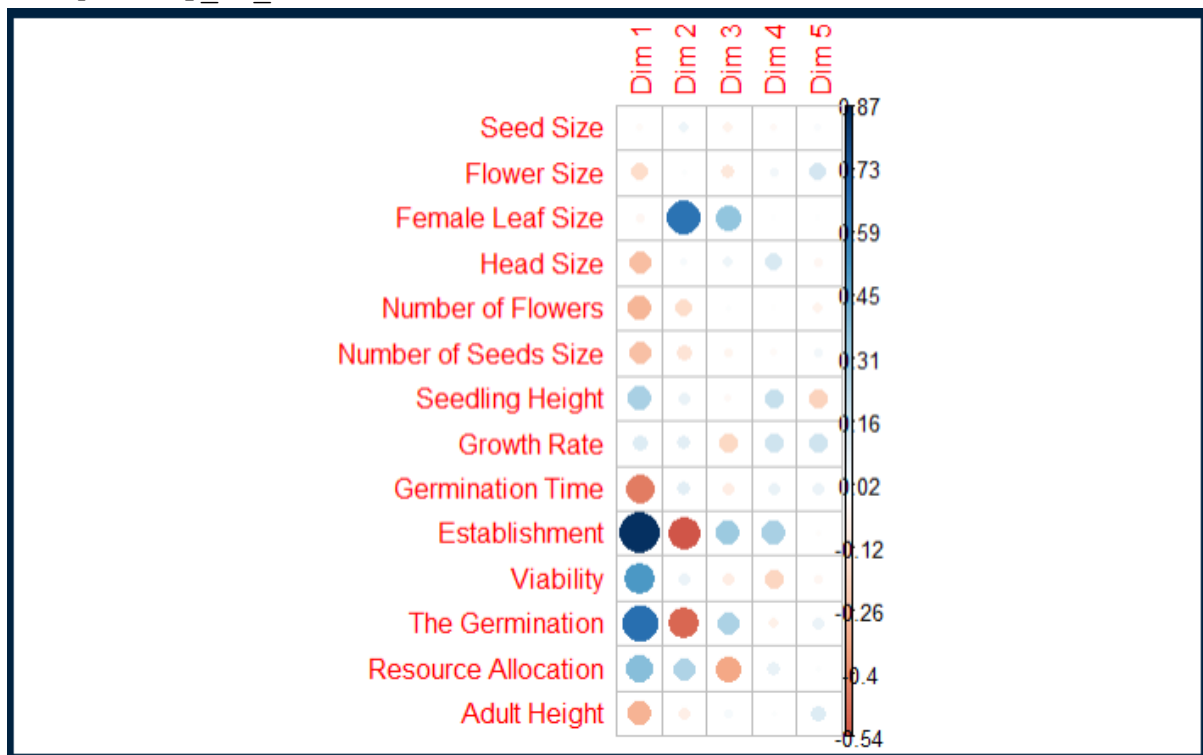
- Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi baris tertinggi hingga terendah. Pada dimensi 1, yang terendah adalah T\_disversifolia dan yang tertinggi adalah T\_tubaeformis.
- Sedangkan untuk kontribusinya, T\_thuberii merupakan yang tertinggi dan T\_longiradiata merupakan yang terendah

```
# Graph of Column variables
my_CA_Col <- my_CA$col
# Coord
fviz_ca_col(my_CA, col.col = "coord",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE)
```



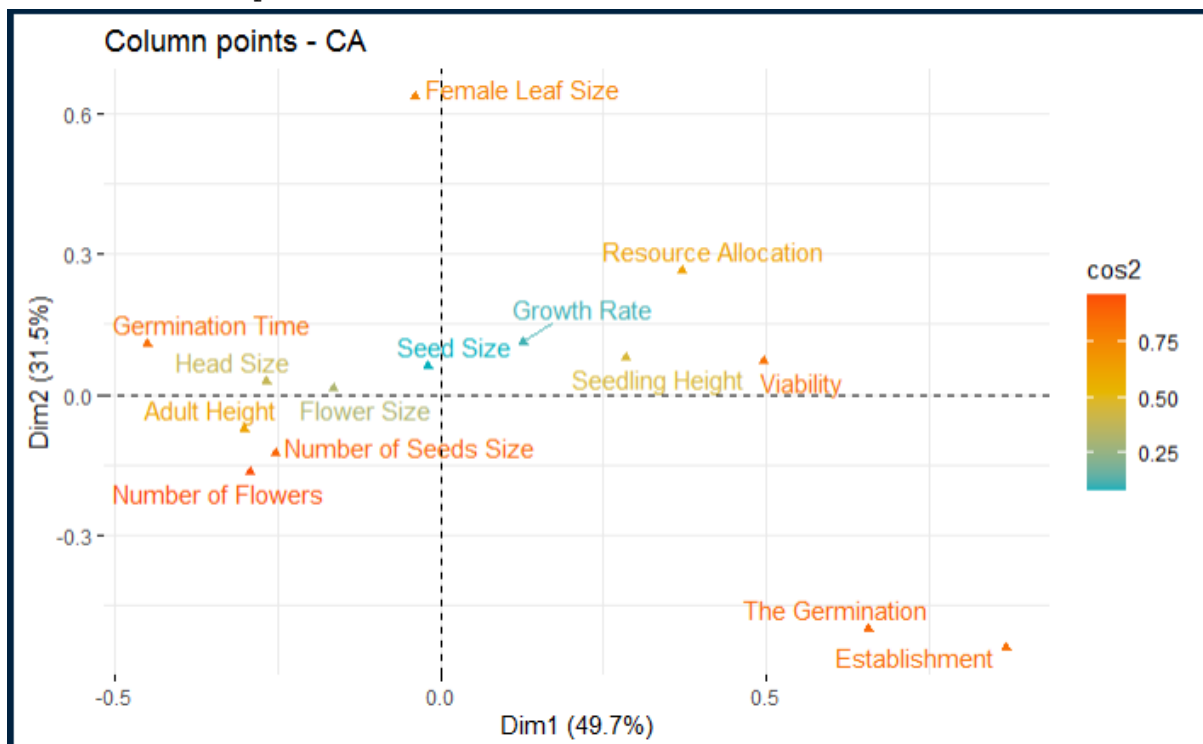
Dari plot CA di atas, data yang telah tersebar adalah 90,2% dan didapatkan bahwa terdapat 2 variabel dengan nilai dimensi 1 positif dan dimensi 2 negatif. Lalu ada 4 variabel dengan nilai kedua dimensi positif. Terdapat 5 variabel dengan nilai dimensi 1 negatif dan dimensi 2 positif. Serta ada 3 variabel dengan nilai kedua dimensi negatif.

```
corrplot(my_CA_Col$coord, is.corr = FALSE)
```



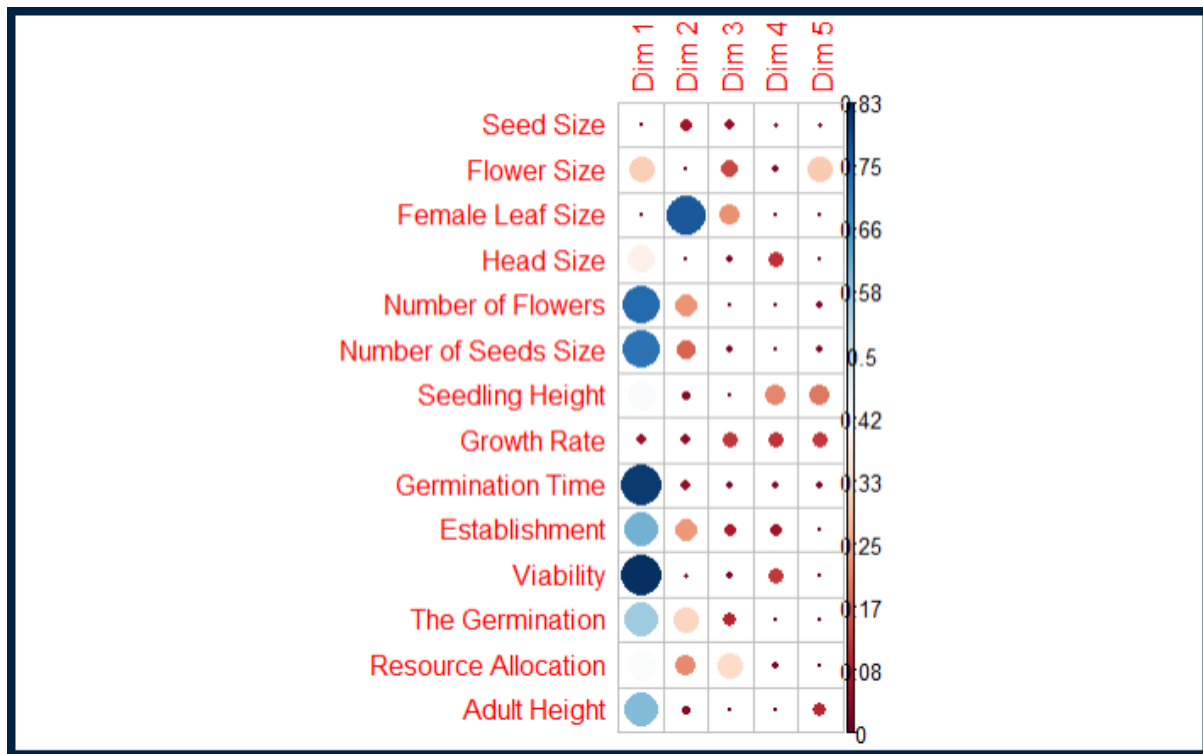
- Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi baris tertinggi hingga terendah. Pada dimensi 1, yang terendah adalah Germination Time dan yang tertinggi adalah Establishment.
- Sedangkan untuk kontribusinya, Establishment merupakan yang tertinggi dan Seed Size merupakan yang terendah

```
# Cos2
fviz_ca_col(my_CA, col.col = "cos2",
            gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
            repel = TRUE)
```



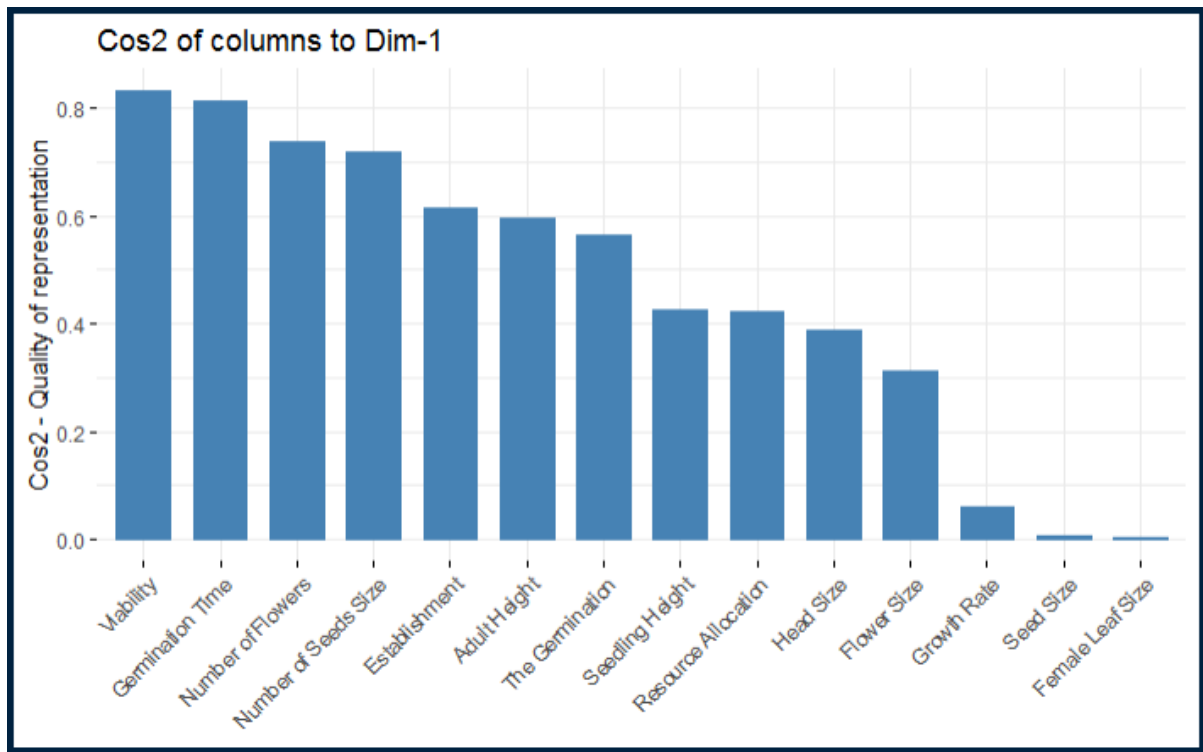
Dari plot CA di atas, data yang telah tersebar adalah 90,2% dan didapati bahwa terdapat 2 variabel dengan nilai dimensi 1 positif dan dimensi 2 negatif. Lalu ada 4 variabel dengan nilai kedua dimensi positif.terdapat 5 variabel dengan nilai dimensi 1 negatif dan dimensi 2 positif. Serta ada 3 variabel dengan nilai kedua dimensi negatif. Yang berbeda dari plot CA sebelumnya adalah skala visualisasi warnanya.

```
corrplot(my_CA_Col$cos2, is.corr = FALSE)
```



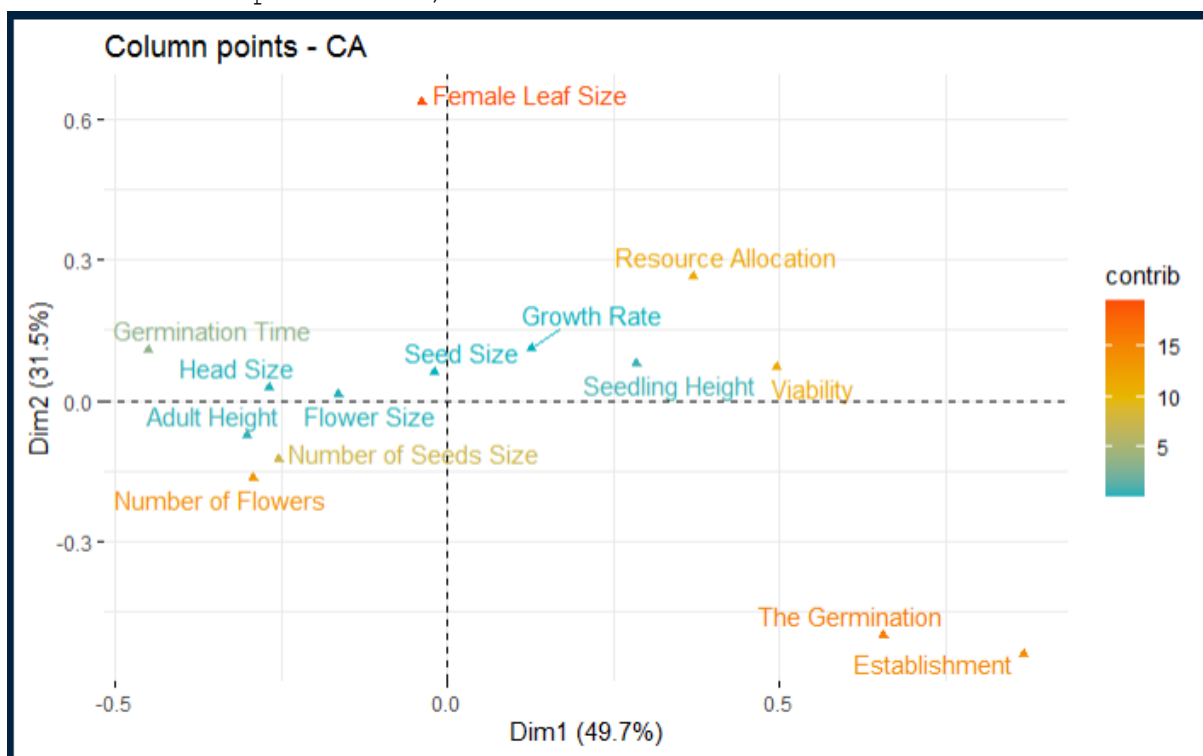
- Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi baris tertinggi hingga terendah. Pada dimensi 1, yang terendah adalah Growth Rate dan yang tertinggi adalah Viability.
- Sedangkan untuk kontribusinya, Establishment merupakan yang tertinggi dan Seed Size merupakan yang terendah

```
fviz_cos2(my_CA, choice = "col")
```



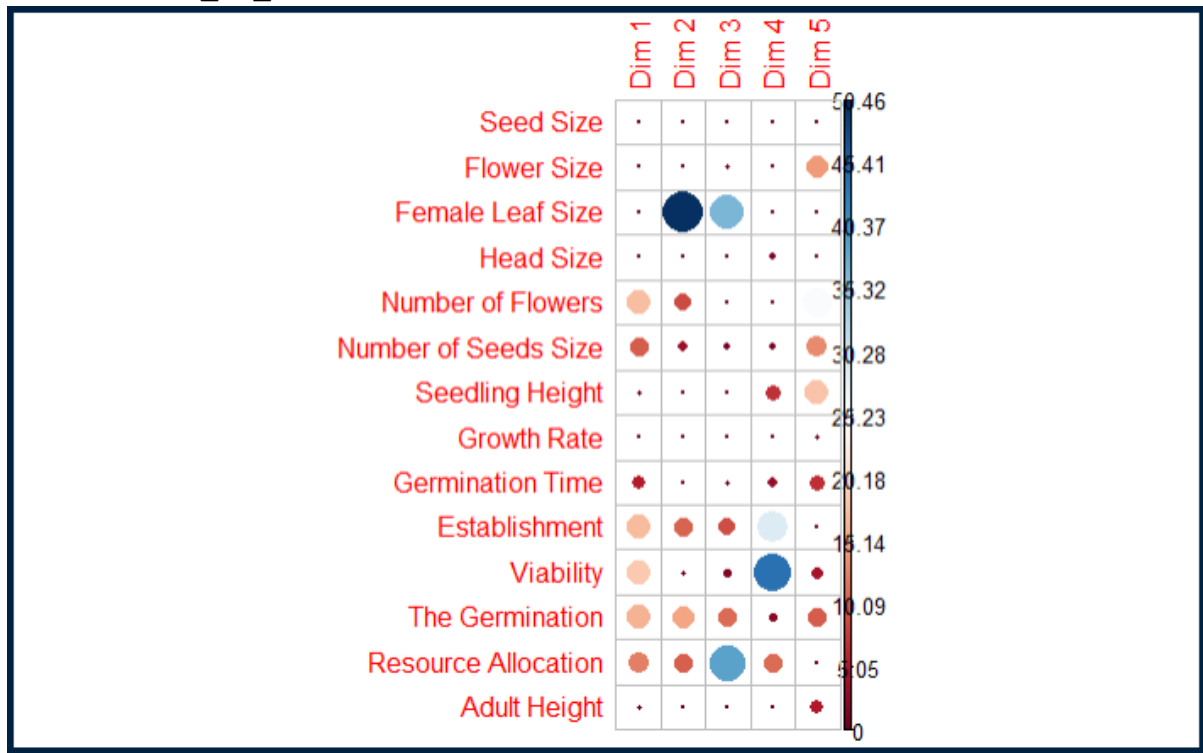
Dari grafik penyebaran dapat kita lihat bahwa tingkat nilai koordinasi kolom tersebar seperti anak tangga. Dan dari plot tersebut terlihat nilai cos2 memiliki penyebaran seperti anak tangga juga.

```
# Contrib
fviz_ca_col(my_CA, col.col = "contrib",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE)
```



Dari plot CA di atas, data yang telah tersebar adalah 90,2% dan didapati bahwa terdapat 2 variabel dengan nilai dimensi 1 positif dan dimensi 2 negatif. Lalu ada 4 variabel dengan nilai kedua dimensi positif. terdapat 5 variabel dengan nilai dimensi 1 negatif dan dimensi 2 positif. Serta ada 3 variabel dengan nilai kedua dimensi negatif. Yang berbeda dari plot CA sebelumnya adalah skala visualisasi warnanya. Pada coord dan cos2, nilai tertinggi adalah 1 sedangkan pada contrib nilai tertinggi adalah 15-20.

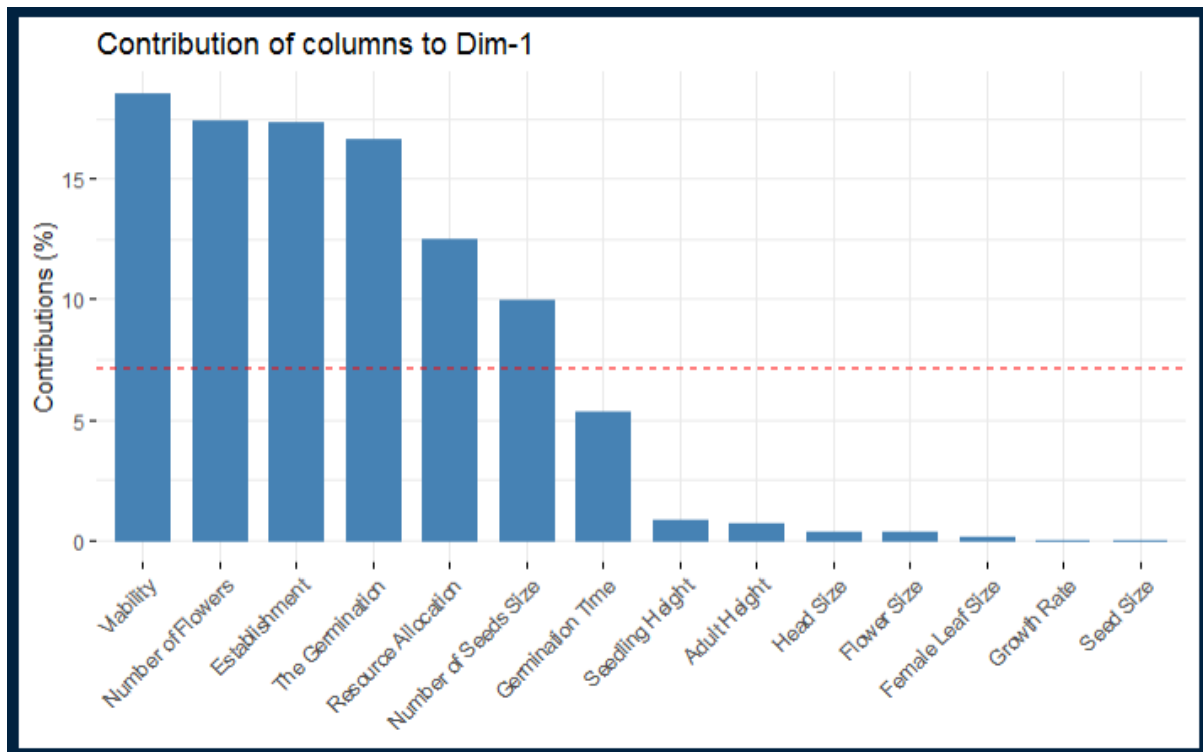
```
corrplot(my_CA_Col$contrib, is.corr = FALSE)
```



- Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi baris tertinggi hingga terendah. Pada dimensi 1, yang terendah adalah Germination Time dan yang tertinggi adalah Viability.
- Sedangkan untuk kontribusinya, The Germination merupakan yang tertinggi dan Seed Size dan Growth Rate merupakan yang terendah

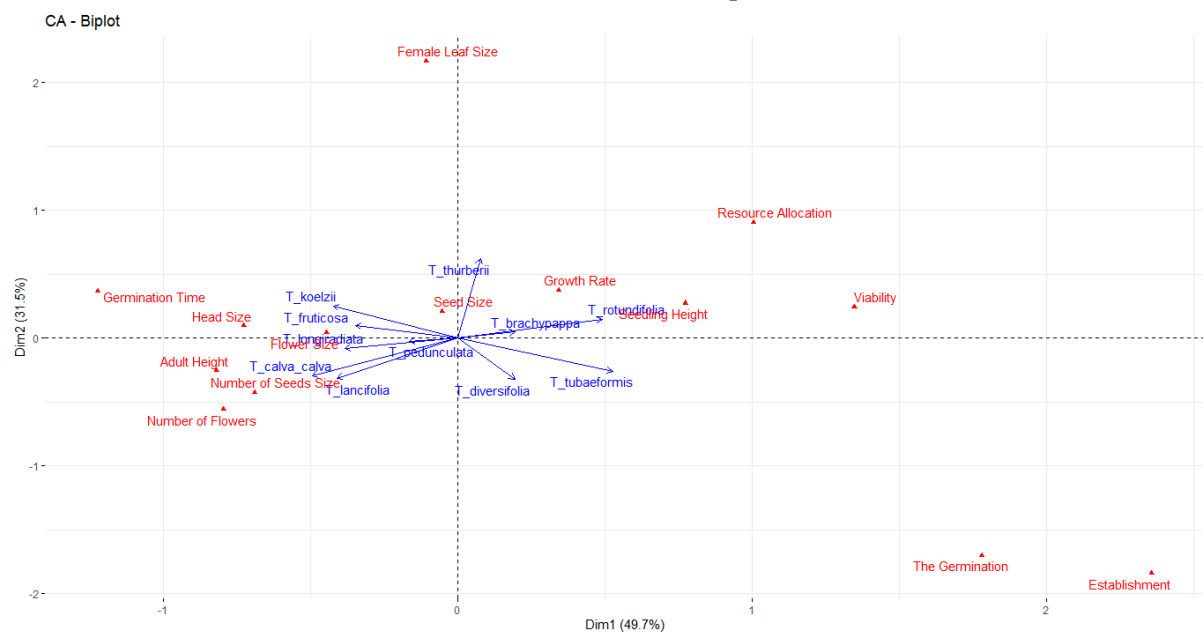
```
fviz_contrib(my_CA, choice = "col")
```





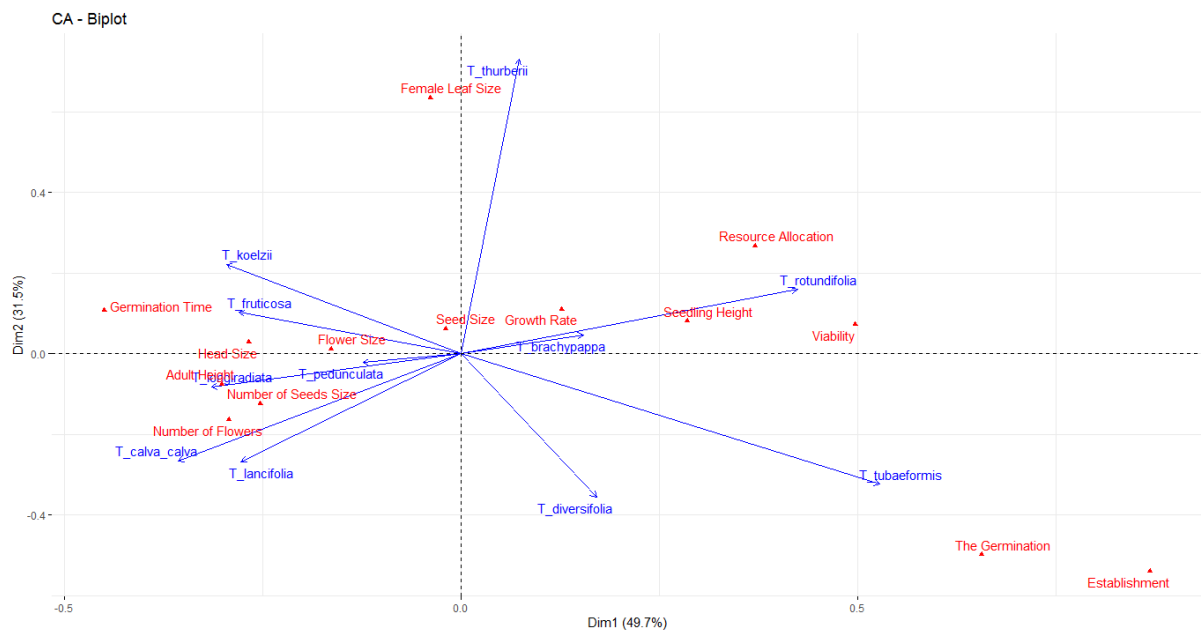
Dari grafik penyebaran dapat kita lihat bahwa tingkat nilai koordinasi kolom tersebar seperti anak tangga. Dan dari plot tersebut terlihat nilai contrib memiliki penyebaran seperti anak tangga juga. 6 variabel pertama memenuhi nilai dalam persebaran (sebagian besar data memiliki nilai yang cukup tinggi).

```
# Biplot Options
# Asymmetric Biplot
fviz_ca_biplot(my_CA, map = "rowprincipal",
               arrow = c(TRUE, FALSE), repel = TRUE)
```



Biplot di atas menunjukkan kesesuaian/dominansi antara variabel dengan spesies. Anak panah yang ditunjukkan oleh garis biru menunjukkan spesies yang menuju pada variabel (merah).

```
# Contribution Biplot
fviz_ca_biplot(my_CA, map = "colgreen",
               arrow = c(TRUE, FALSE), repel = TRUE)
```



Biplot di atas menunjukkan kesesuaian/dominansi antara variabel dengan spesies. Anak panah yang ditunjukkan oleh garis biru menunjukkan spesies yang paling banyak berkontribusi (bernilai besar) pada variabel (merah).

```
# Dimension Description
CA_Desc <- dimdesc(my_CA)
CA_Desc
``,``
```

\$'Dim 1'		\$'Dim 2'		\$'Dim 3'	
\$'Dim 1'\$row		\$'Dim 2'\$row		\$'Dim 3'\$row	
	coord		coord		coord
T_calva_calva	-0.49359664	T_diversifolia	-0.32225871	T_brachypappa	-0.45828421
T_koelzii	-0.42244539	T_lancifolia	-0.31509924	T_pedunculata	-0.21271227
T_lancifolia	-0.40964534	T_calva_calva	-0.29295516	T_fruticosa	-0.07636381
T_longiradiata	-0.38181579	T_tubaeformis	-0.25766462	T_calva_calva	-0.05592861
T_fruticosa	-0.34681549	T_longiradiata	-0.07941980	T_lancifolia	-0.04165190
T_pedunculata	-0.16365989	T_pedunculata	-0.02318814	T_rotundifolia	-0.03497129
T_thurberii	0.07790628	T_brachypappa	0.04731343	T_tubaeformis	0.06593261
T_diversifolia	0.19439202	T_fruticosa	0.10062935	T_longiradiata	0.09110935
T_brachypappa	0.19517260	T_rotundifolia	0.14716143	T_thurberii	0.11212118
T_rotundifolia	0.49291522	T_koelzii	0.25139938	T_koelzii	0.22947648
T_tubaeformis	0.52797690	T_thurberii	0.62152377	T_diversifolia	0.25121287
\$'Dim 1'\$col		\$'Dim 2'\$col		\$'Dim 3'\$col	
	coord		coord		coord
Germination Time	-0.44960618	Establishment	-0.53890783	Resource Allocation	-0.33190883
Adult Height	-0.30118745	The Germination	-0.49816238	Growth Rate	-0.17702472
Number of Flowers	-0.29248590	Number of Flowers	-0.16350142	Flower Size	-0.10968081
Head Size	-0.26738451	Number of Seeds Size	-0.12390191	Viability	-0.08178190
Number of Seeds Size	-0.25333237	Adult Height	-0.07519333	Germination Time	-0.08102201
Flower Size	-0.16372654	Flower Size	0.01216981	Seed Size	-0.05374186
Female Leaf Size	-0.03879293	Head Size	0.02791394	Number of Seeds Size	-0.04517486
Seed Size	-0.01904995	Seed Size	0.06153008	Seedling Height	-0.03004661
Growth Rate	0.12691279	Viability	0.07210812	Number of Flowers	0.01004261
Seedling Height	0.28514978	Seedling Height	0.08030535	Adult Height	0.03300887
Resource Allocation	0.37060014	Germination Time	0.10759097	Head Size	0.06080137
Viability	0.49640774	Growth Rate	0.11034971	The Germination	0.27082735
The Germination	0.65616089	Resource Allocation	0.26602933	Establishment	0.31769222
Establishment	0.86834684	Female Leaf Size	0.63472287	Female Leaf Size	0.34594890

## Kesimpulan

Data *Tihtonia* dapat diolah menggunakan metode CA yang didasarkan pada variabel dan individual dari data. Data tersebut diolah sehingga dapat menghasilkan individual yang memiliki variabel (karakter) mencolok. Dan dari beberapa metode (coord, cos2, contrib) juga terdapat berbagai hasil.

```

```{R}
# MCA
# Dataset Preparation
data("quine")
head(quine)

```

Description: `df[,5] [6 x 5]`

	Eth <fctr>	Sex <fctr>	Age <fctr>	Lrn <fctr>	Days <int>
1	A	M	F0	SL	2
2	A	M	F0	SL	11
3	A	M	F0	SL	14
4	A	M	F0	AL	5
5	A	M	F0	AL	5
6	A	M	F0	AL	13

6 rows

```

Data_MCA <- MCA(quine[, 1:4], graph = FALSE)
Data_MCA

```

```

**Results of the Multiple Correspondence Analysis (MCA)**
The analysis was performed on 146 individuals, described by 4 variables
*The results are available in the following objects:

  name          description
1  "$eig"       "eigenvalues"
2  "$var"       "results for the variables"
3  "$var$coord" "coord. of the categories"
4  "$var$cos2"  "cos2 for the categories"
5  "$var$contrib" "contributions of the categories"
6  "$var$v.test" "v-test for the categories"
7  "$ind"       "results for the individuals"
8  "$ind$coord" "coord. for the individuals"
9  "$ind$cos2"  "cos2 for the individuals"
10 "$ind$contrib" "contributions of the individuals"
11 "$call"      "intermediate results"
12 "$call$marge.col" "weights of columns"
13 "$call$marge.li" "weights of rows"

```

MCA digunakan untuk menganalisis kumpulan data dari survei. Tujuannya adalah untuk mengidentifikasi Sekelompok individu dengan profil yang sama. Dari output tersebut, bahwasanya Analisis dilakukan pada 146 individu, dijelaskan oleh 4 variabel.

Hasilnya tersedia di 13 objek

```

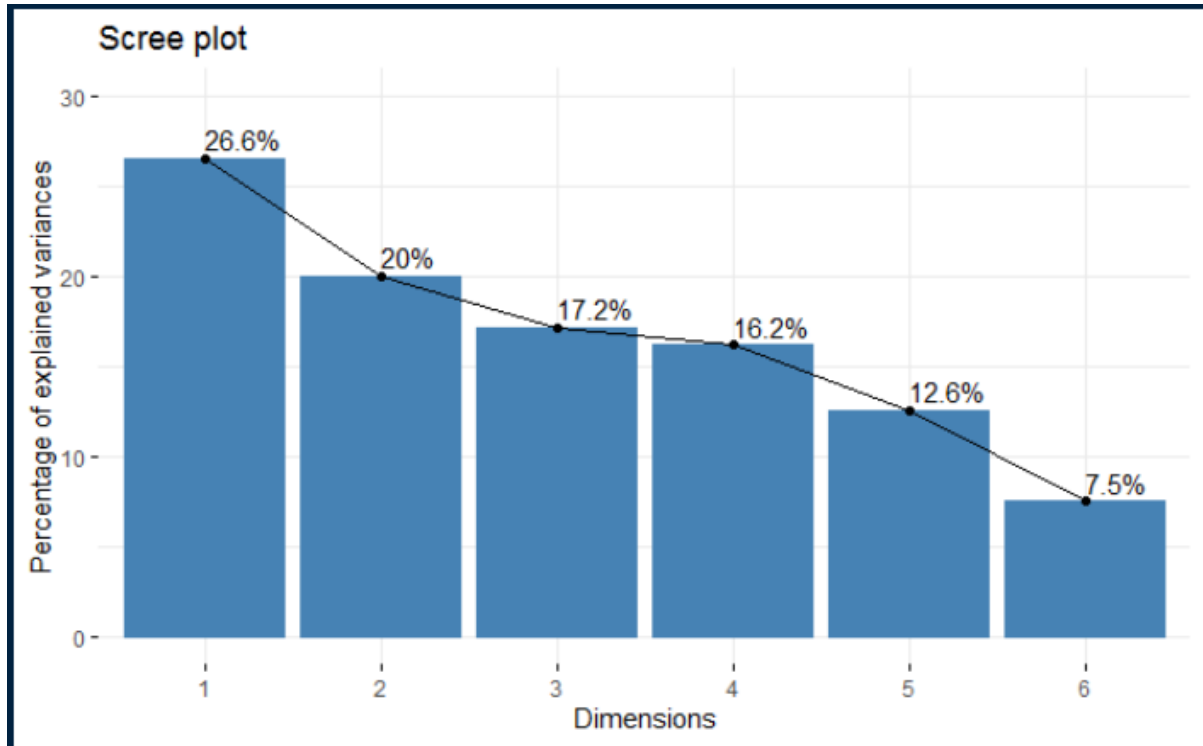
# Eigenvalues / Variances
Data_MCA$eig

```

	eigenvalue	percentage of variance	cumulative percentage of variance
dim 1	0.3982765	26.551764	26.55176
dim 2	0.2996461	19.976406	46.52817
dim 3	0.2577774	17.185163	63.71333
dim 4	0.2426799	16.178660	79.89199
dim 5	0.1886977	12.579843	92.47184
dim 6	0.1129225	7.528164	100.00000

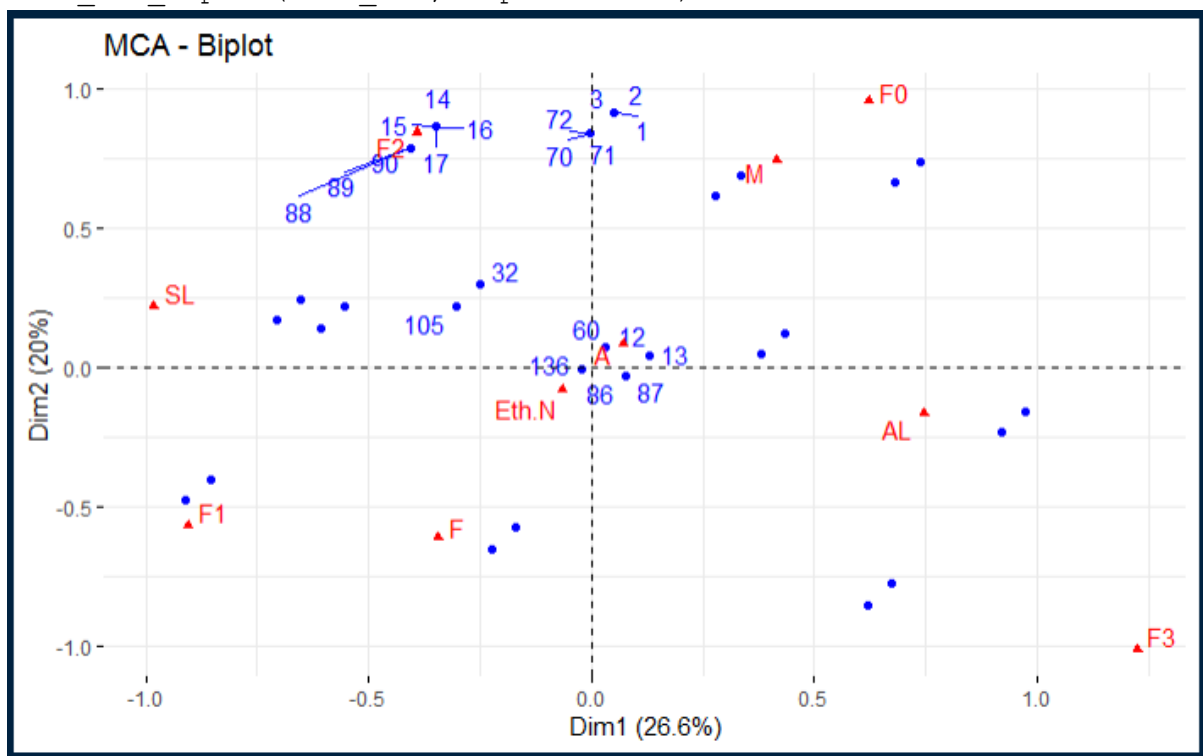
Didapatkan dari perhitungan tersebut didapat 6 nilai eigen yang diurutkan dari varians terbesar yaitu 26.551 sehingga pada dimensi 1 nilai eigennya adalah 0.392765 dan seterusnya.

```
fviz_screplot(Data_MCA, addlabels = TRUE, ylim = c(0, 30))
```



Screeplot di atas menunjukkan presentase nilai variance pada tiap dimensi. Dimensi 1 nilainya adalah 26,6% dan seterusnya.

```
# Biplot
fviz_mca_biplot(Data_MCA, repel = TRUE)
```



Pada biplot di atas, hanya ada 46,6% persebaran data. Faktor pada setiap variabel dijelaskan, sehingga variabelnya menjadi lebih banyak. AL (Lrn/status learner) dan F3 (age) memiliki

nilai dimensi 1 positif dan dimensi 2 negatif dengan kurang lebih 4 titik dominan. A (Eth/ethnic), M (sex) dan F0 (age) memiliki nilai dimensi 1 dan 2 yang positif dengan kurang lebih 8 titik dominan. SL (Lrn) dan F2 (age) memiliki nilai dimensi 2 positif dan dimensi 1 negatif dengan kurang lebih 9 titik dominan. N (Eth), F (sex), dan F1 (age) memiliki nilai dimensi 1 dan dimensi 2 yang negative dengan kurang lebih 4 titik dominan. Setiap titik biru pada biplot di atas memungkinkan berisi banyak data.

```
# Graph of Variables
Data_MCA$var
```

\$coord					
	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
A	0.07288947	0.08631111	-0.75968509	0.71410580	-0.126890216
Eth.N	-0.06531654	-0.07734372	0.68075677	-0.63991299	0.113706817
F	-0.34404494	-0.61265706	-0.10432058	0.10350590	0.550753959
M	0.41702417	0.74261462	0.12644919	-0.12546169	-0.667580557
F0	0.62413540	0.95778656	1.03510412	1.10389575	0.881783732
F1	-0.90507992	-0.57054472	0.42424557	0.36821316	-0.754473540
F2	-0.39125161	0.84617330	-0.88816163	-0.85337512	0.400799432
F3	1.22521470	-1.01400340	-0.36171643	-0.38206018	-0.155586521
AL	0.74615671	-0.16567598	0.01280462	-0.01176222	0.007212796
SL	-0.98303185	0.21827153	-0.01686958	0.01549625	-0.009502573

\$contrib					
	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
A	0.1576091	0.2937386	26.452001104	24.827189843	1.008152018
Eth.N	0.1412341	0.2632203	23.703741249	22.247741548	0.903408951
F	4.0712039	17.1594519	0.578325379	0.604746744	22.020435370
M	4.9347926	20.7993357	0.701000459	0.733026357	26.691436813
F0	4.5219308	14.1540189	19.216490676	23.215243519	19.050596872
F1	16.2007159	8.5568860	5.499645591	4.400578728	23.761098613
F2	2.6325388	16.3665509	20.959757235	20.553855501	5.830889948
F3	21.2980994	19.3897091	2.868089174	3.398840934	0.724900735
AL	19.8673846	1.3018943	0.009039695	0.008102329	0.003918376
SL	26.1744908	1.7151941	0.011909439	0.010674497	0.005162304

\$cos2					
	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
A	0.004760888	0.006675623	0.5171607635	0.4569655757	1.442828e-02
Eth.N	0.004760888	0.006675623	0.5171607635	0.4569655757	1.442828e-02
F	0.143475057	0.454968096	0.0131912538	0.0129860254	3.676726e-01
M	0.143475057	0.454968096	0.0131912538	0.0129860254	3.676726e-01
F0	0.088384160	0.208139390	0.2430999553	0.2764858610	1.764172e-01
F1	0.376818048	0.149739788	0.0827927798	0.0623672278	2.618459e-01
F2	0.057765216	0.270192174	0.2976721086	0.2748109786	6.061894e-02
F3	0.438389248	0.300271640	0.0382095539	0.0426284009	7.069349e-03
AL	0.733495812	0.036162351	0.0002160086	0.0001822703	6.854012e-05
SL	0.733495812	0.036162351	0.0002160086	0.0001822703	6.854012e-05

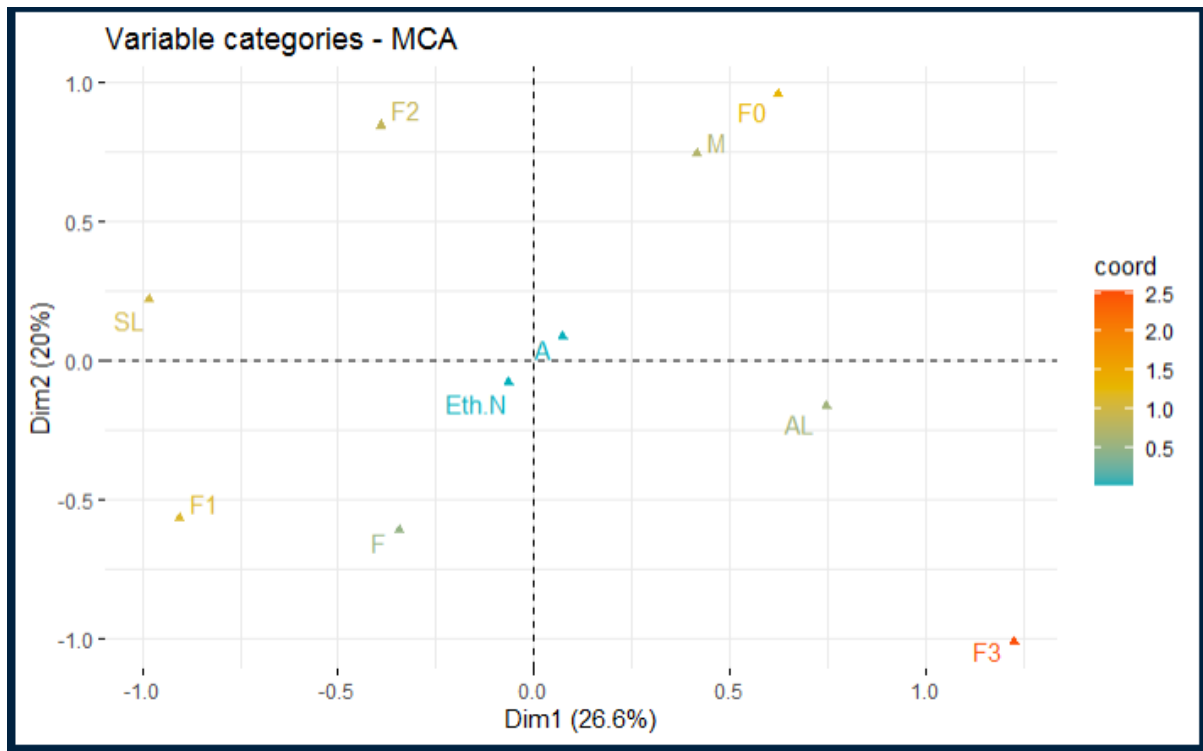
\$v.test					
	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
A	0.8308602	0.9838523	-8.6595791	8.1400251	-1.44640968
Eth.N	-0.8308602	-0.9838523	8.6595791	-8.1400251	1.44640968
F	-4.5611274	-8.1222148	-1.3830155	1.3722149	7.30154313
M	4.5611274	8.1222148	1.3830155	-1.3722149	-7.30154313
F0	3.5799027	5.4936519	5.9371284	6.3317020	5.05771653
F1	-7.3917939	-4.6596426	3.4648165	3.0071994	-6.16179053
F2	-2.8941245	6.2592224	-6.5698140	-6.3124949	2.96475057
F3	7.9728565	-6.5984383	-2.3538023	-2.4861855	-1.01245032
AL	10.3129478	-2.2898779	0.1769781	-0.1625706	0.09969111
SL	-10.3129478	2.2898779	-0.1769781	0.1625706	-0.09969111

\$eta2					
	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
Eth	0.004760888	0.006675623	0.5171607635	0.4569655757	1.442828e-02
Sex	0.143475057	0.454968096	0.0131912538	0.0129860254	3.676726e-01
Age	0.711374095	0.700778301	0.5005417518	0.5005857074	3.726211e-01
Lrn	0.733495812	0.036162351	0.0002160086	0.0001822703	6.854012e-05

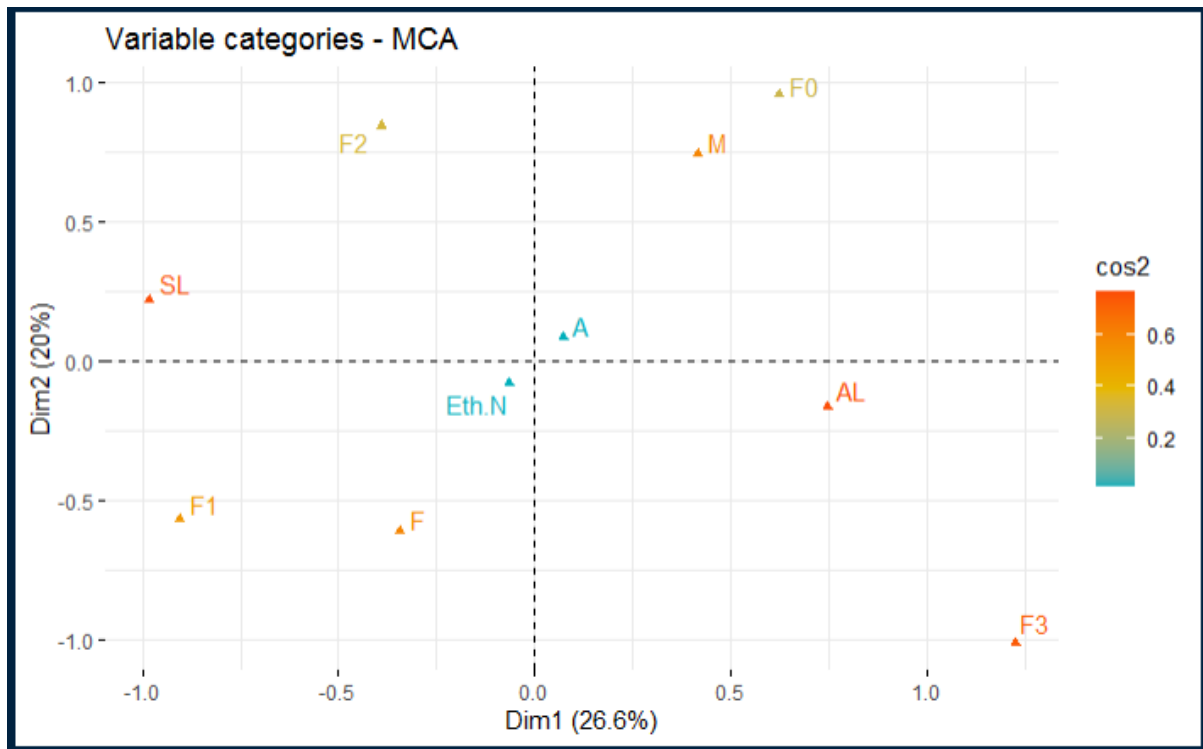
\$coord merupakan nilai dari suatu koordinat untuk kategori  
 \$cos2 memiliki nilai untuk kualitas representasi kategori  
 \$contrib adalah nilai kontribusi kategori dalam persen untuk definisi dimensi

```
# Coord
fviz_mca_var(Data_MCA, col.var = "coord",
              gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
              repel = TRUE)
```



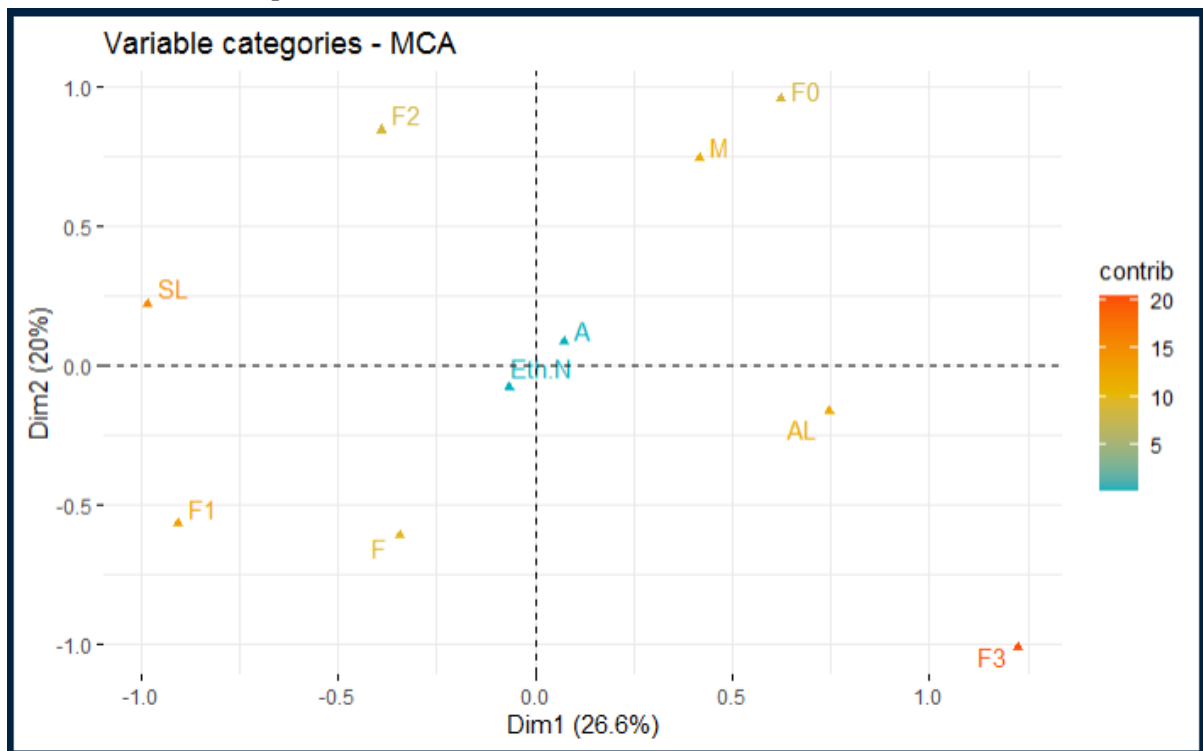
Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi kategori terendah hingga tertinggi. Yang tertinggi adalah kategori umur F3 dan yang terendah adalah kategory etnis/suku bangsa.

```
# Cos2
fviz_mca_var(Data_MCA, col.var = "cos2",
              gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
              repel = TRUE)
```



Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi kategori terendah hingga tertinggi. Yang tertinggi adalah kategori status learner dan yang terendah adalah kategory etnis/suku bangsa.

```
# Contrib
fviz_mca_var(Data_MCA, col.var = "contrib",
              gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
              repel = TRUE)
```





Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi kategori terendah hingga tertinggi. Yang tertinggi adalah kategori umur F3 dan yang terendah adalah kategory etnis/suku bangsa.

```
# Graph of Individuals
```

```
Data_MCA$ind
```

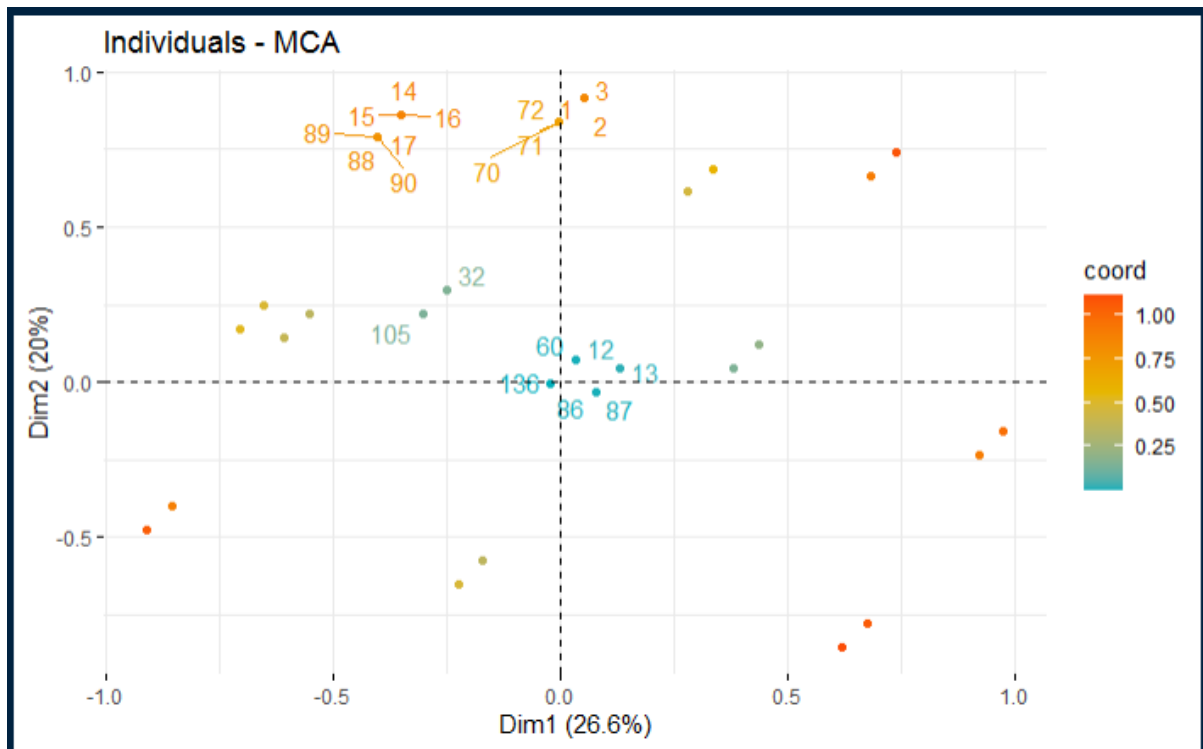
\$coord	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
1	0.051901029	0.91568600	0.18957312	0.86680251	0.04478106
2	0.051901029	0.91568600	0.18957312	0.86680251	0.04478106
3	0.051901029	0.91568600	0.18957312	0.86680251	0.04478106
4	0.736900195	0.74033527	0.20418468	0.85296924	0.05440100
5	0.736900195	0.74033527	0.20418468	0.85296924	0.05440100
6	0.736900195	0.74033527	0.20418468	0.85296924	0.05440100
7	0.736900195	0.74033527	0.20418468	0.85296924	0.05440100
8	0.736900195	0.74033527	0.20418468	0.85296924	0.05440100
9	-0.553880923	0.21768957	-0.11121329	0.49345471	-0.89690979
10	-0.553880923	0.21768957	-0.11121329	0.49345471	-0.89690979

\$contrib	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
1	4.632490e-03	1.916600826	0.095489528	2.1205751233	0.007278960
2	4.632490e-03	1.916600826	0.095489528	2.1205751233	0.007278960
3	4.632490e-03	1.916600826	0.095489528	2.1205751233	0.007278960
4	9.338559e-01	1.252839419	0.110776729	2.0534308806	0.010742231
5	9.338559e-01	1.252839419	0.110776729	2.0534308806	0.010742231
6	9.338559e-01	1.252839419	0.110776729	2.0534308806	0.010742231
7	9.338559e-01	1.252839419	0.110776729	2.0534308806	0.010742231
8	9.338559e-01	1.252839419	0.110776729	2.0534308806	0.010742231
9	5.275885e-01	0.108321275	0.032863636	0.6872392333	2.919968614
10	5.275885e-01	0.108321275	0.032863636	0.6872392333	2.919968614

\$cos2	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
1	1.338006e-03	4.164848e-01	0.0178508759	3.732040e-01	0.0009960811
2	1.338006e-03	4.164848e-01	0.0178508759	3.732040e-01	0.0009960811
3	1.338006e-03	4.164848e-01	0.0178508759	3.732040e-01	0.0009960811
4	2.898240e-01	2.925323e-01	0.0222517022	3.883146e-01	0.0015795405
5	2.898240e-01	2.925323e-01	0.0222517022	3.883146e-01	0.0015795405
6	2.898240e-01	2.925323e-01	0.0222517022	3.883146e-01	0.0015795405
7	2.898240e-01	2.925323e-01	0.0222517022	3.883146e-01	0.0015795405
8	2.898240e-01	2.925323e-01	0.0222517022	3.883146e-01	0.0015795405
9	2.108686e-01	3.257274e-02	0.0085014387	1.673685e-01	0.5529381863
10	2.108686e-01	3.257274e-02	0.0085014387	1.673685e-01	0.5529381863

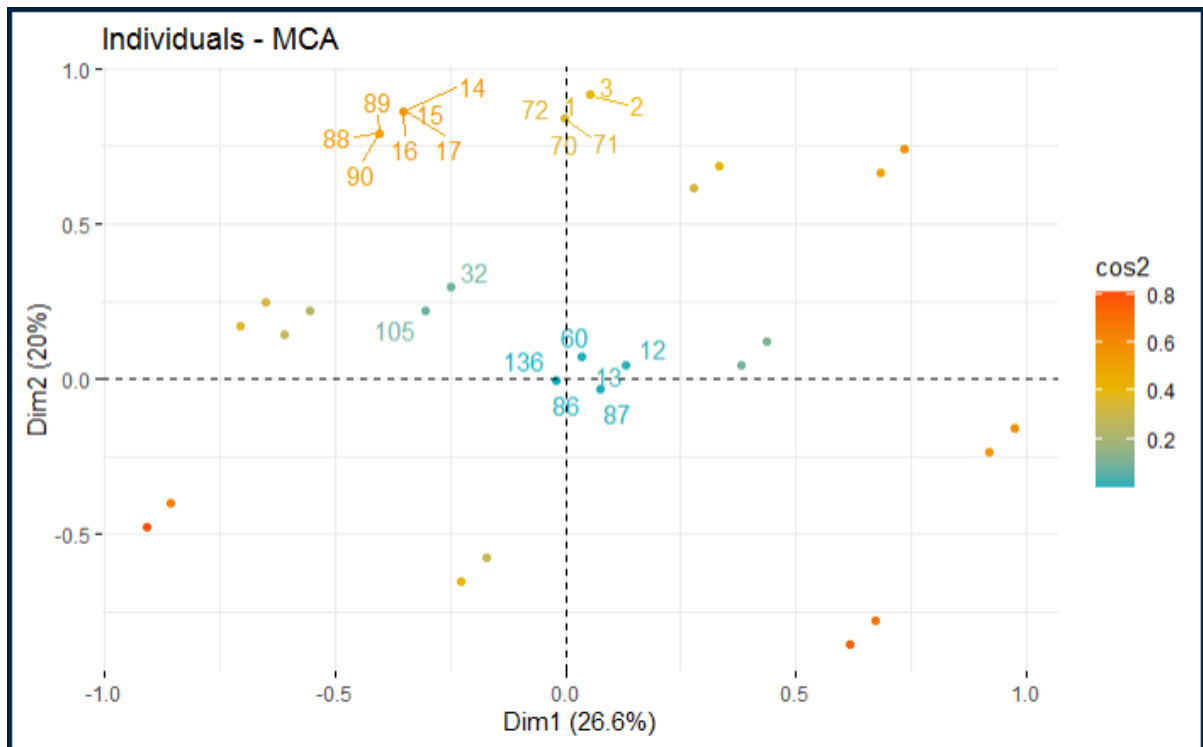
```
# Coord
```

```
fviz_mca_ind(Data_MCA, col.ind = "coord",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE)
```



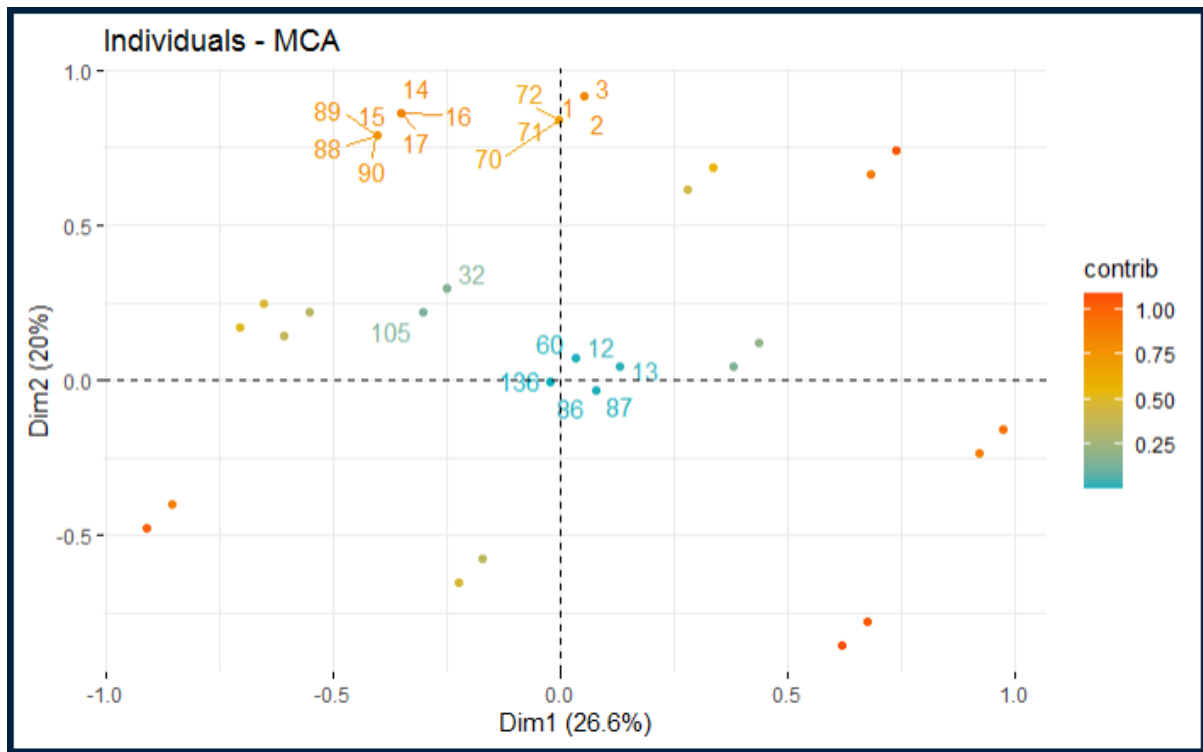
Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi kategori terendah hingga tertinggi. Yang tertinggi berada jauh dari pusat dimensi, dan yang terendah berada di sekitar pusat dimensi seperti urutan ke-60, 12, 13, 87, dan lainnya.

```
# Cos2
fviz_mca_ind(Data_MCA, col.ind = "cos2",
              gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
              repel = TRUE)
```



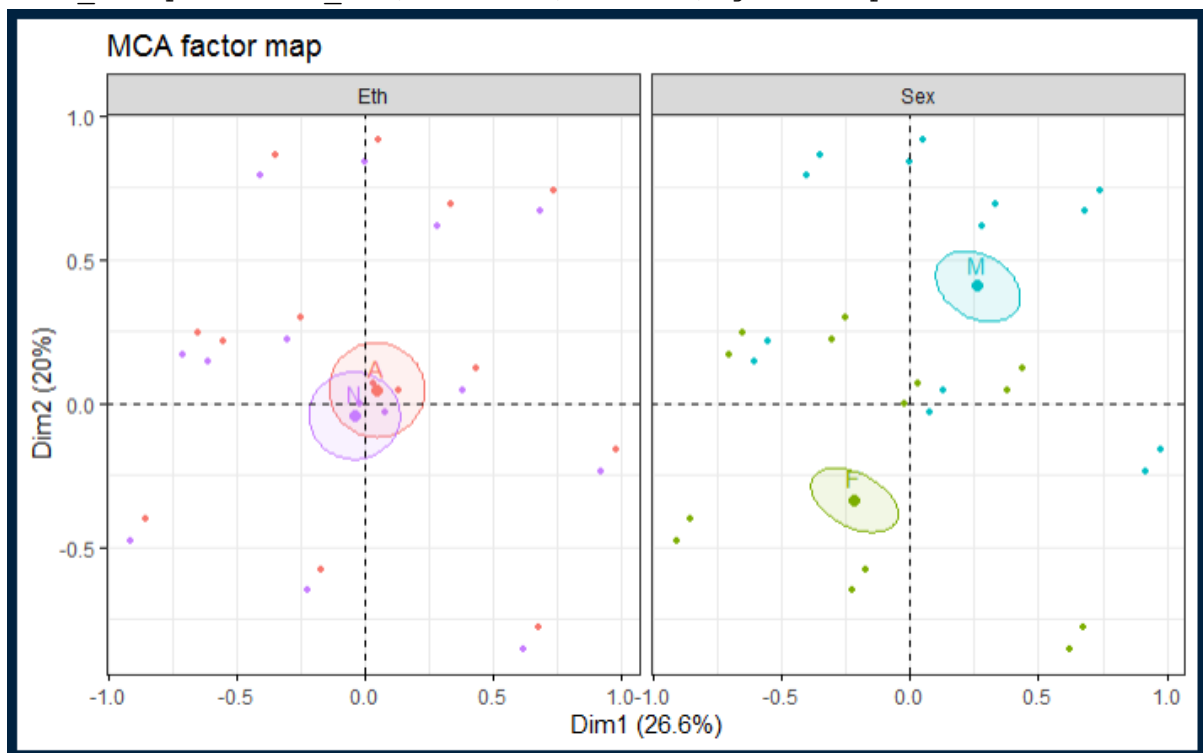
Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai koordinasi kategori terendah hingga tertinggi. Yang tertinggi berada jauh dari pusat dimensi, dan yang terendah berada di sekitar pusat dimensi seperti urutan ke-60, 12, 13, 87, dan lainnya.

```
# Contrib
fviz_mca_ind(Data_MCA, col.ind = "contrib",
              gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
              repel = TRUE)
```



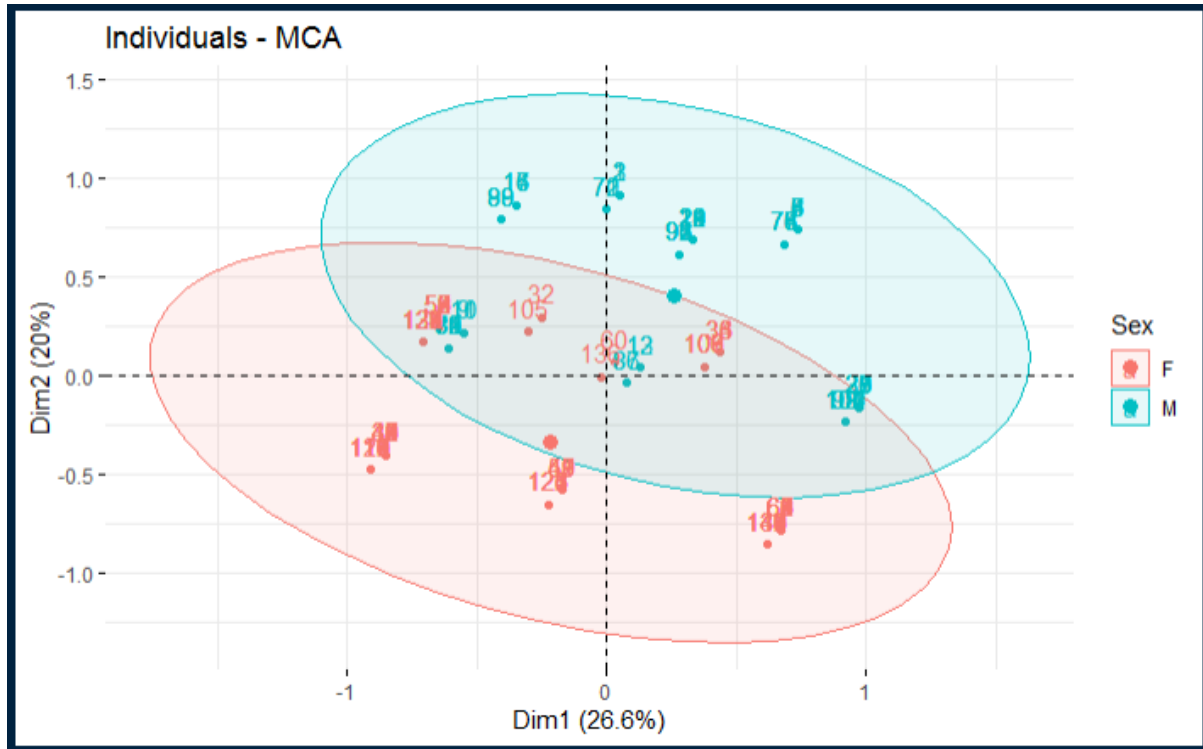
Dari grafik penyebaran dapat kita lihat bahwa gradasi warna yang dimulai dari biru hingga kemerahan menandakan tingkat nilai kontribusi kategori terendah hingga tertinggi. Yang tertinggi berada jauh dari pusat dimensi, dan yang terendah berada di sekitar pusat dimensi seperti urutan ke-60, 12, 13, 87, dan lainnya.

```
# Color Individuals by Groups
fviz_ellipses(Data_MCA, c("Eth", "Sex"), geom = "point")
```



Dari plot di atas, dapat diperoleh informasi mengenai pewarnaan data sesuai dengan faktornya. Pada variabel Eth terdapat 2 faktor yaitu A (Aborigin) dan N (Non-Aborigin). Lalu pada variabel sex terdapat 2 faktor juga yaitu M (Male) dan F (Female)

```
fviz_mca_ind(Data_MCA, habillage = 2, addEllipses = TRUE)
```

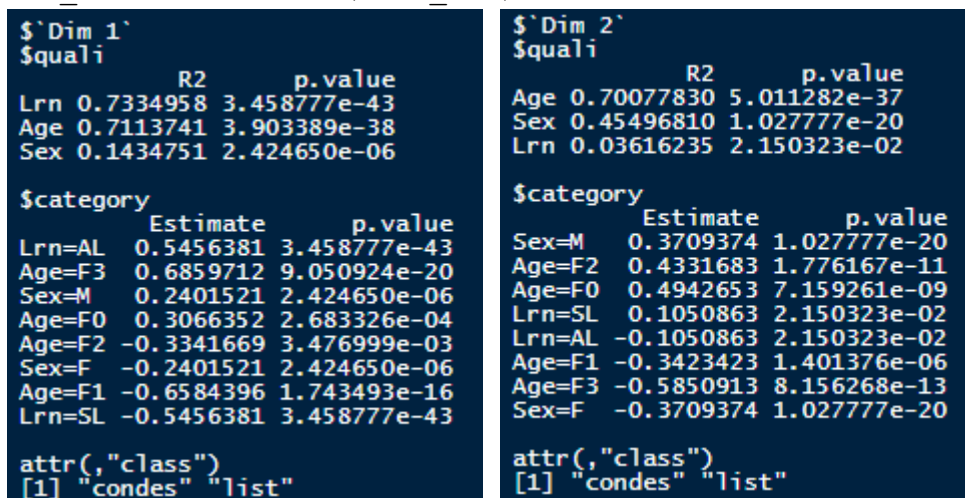


Berdasarkan plot diatas mengenai jenis kelaminnya, terbagi menjadi 2 kelompok yaitu Male dan Female. Dan dapat dilihat dari plot tersebut keduanya memiliki kelompok penyebaran yang hamper sama banyaknya.

```
# Habillage = Index of The Column to be Used as Grouping variable
```

```
# dimension Description
```

```
MCA_Desc <- dimdesc(Data_MCA)
```



MCA\_Desc  
...

Description: df[,5] [146 x 5]					
	Dim 1 <dbl>	Eth <fctr>	Sex <fctr>	Age <fctr>	Ln <fctr>
1	0.051901029	Eth=A	Sex=M	Age=F0	Ln=SL
2	0.051901029	Eth=A	Sex=M	Age=F0	Ln=SL
3	0.051901029	Eth=A	Sex=M	Age=F0	Ln=SL
4	0.736900195	Eth=A	Sex=M	Age=F0	Ln=AL
5	0.736900195	Eth=A	Sex=M	Age=F0	Ln=AL
6	0.736900195	Eth=A	Sex=M	Age=F0	Ln=AL
7	0.736900195	Eth=A	Sex=M	Age=F0	Ln=AL
8	0.736900195	Eth=A	Sex=M	Age=F0	Ln=AL
9	-0.553880923	Eth=A	Sex=M	Age=F1	Ln=SL
10	-0.553880923	Eth=A	Sex=M	Age=F1	Ln=SL
11	-0.553880923	Eth=A	Sex=M	Age=F1	Ln=SL
12	0.131118242	Eth=A	Sex=M	Age=F1	Ln=AL
13	0.131118242	Eth=A	Sex=M	Age=F1	Ln=AL
14	-0.350333448	Eth=A	Sex=M	Age=F2	Ln=SL
15	-0.350333448	Eth=A	Sex=M	Age=F2	Ln=SL
16	-0.350333448	Eth=A	Sex=M	Age=F2	Ln=SL
17	-0.350333448	Eth=A	Sex=M	Age=F2	Ln=SL
18	0.334665717	Eth=A	Sex=M	Age=F2	Ln=AL
19	0.334665717	Eth=A	Sex=M	Age=F2	Ln=AL
20	0.334665717	Eth=A	Sex=M	Age=F2	Ln=AL
21	0.334665717	Eth=A	Sex=M	Age=F2	Ln=AL

1-21 of 146 rows

Previous1234567Next

**Kesimpulan**

Data *quine* yang diolah menggunakan MCA menghasilkan plot yang membaginya berdasarkan faktor dari variabel. Data tersebut juga dapat diketahui memiliki kesamaan antar isinya. Selain itu, penggunaan MCA juga dapat membuat data lebih ringkas untuk divisualisasikan.